

Mad cows and the minister

A realistic assessment of the risk to human health from bovine spongiform encephalopathy, rather than the hollow pronouncement that "beef is safe" would best counter public fears in Britain.

PUBLIC alarm, even hysteria, has been reawakened in Britain over the possible consequences for human health of the outbreak of bovine spongiform encephalopathy (BSE) in the cattle herd. Several local education authorities have decided to ban British beef from school kitchens. And there are signs that general demand for beef has fallen — to the chagrin of breeders and butchers, wholesale beef prices are tumbling. Agriculture minister John Gummer is under great pressure to do more to protect public health; his difficulty is that, while huge uncertainties persist about the transfer of the BSE agent between species, people included, he is compelled by the nature of his office to provide public reassurance that must seem hollow. His hand may now be forced by the House of Commons agriculture select committee (which savaged the government last year over its handling of a salmonella outbreak in British eggs in 1988).

The origins of the scare over BSE are not far to seek. BSE is the close analogue in cattle of the disease of sheep called scrapie, caused by an infectious agent which is neither a virus nor a bacterium, and possibly a mere polypeptide called a prion. (Those offended by the notion that protein molecules may influence whole cells should concede that regulatory proteins might do just that.) The report a year ago, by a committee under Sir Richard Southwood, of the Department of Zoology at the University of Oxford, suggested that BSE may have arisen in British cattle through the practice, in the early 1980s, of feeding processed sheep offal to cattle, but the evidence for even that inference is circumstantial. Although scrapie is spread in sheep by maternal transmission, there is no evidence of that in cattle. The Southwood committee remarked on the lack of evidence, but was careful not to exclude the possibility. What does seem reasonably certain is that maternal transfer is the only mechanism of transmission of scrapie in sheep, and that in both sheep and cattle, the infection mostly resides in nervous tissue.

People in Britain are not worried about the welfare of cattle, but about themselves and their children. If the scrapie agent has transferred from sheep to cattle, will it now infect people? The discovery that at least one cat in Britain has died of an analogous disease (see *Nature* 345, 194; 17 May 1990) has triggered the latest alarm by putting in people's minds the notion that inter-specific transfer may be general. If more cats, or animals of other species, die of similar causes, the sense of alarm will

certainly be strengthened.

Two difficulties arise. If transmission from cattle to people is possible, and has happened, the disease that will result would probably be indistinguishable from Creutzfeldt-Jakob disease (CJD), with symptoms of dementia and characteristic post mortem lesions in the brain. That is why the British agriculture ministry has sensibly commissioned research on the epidemiology of that disease. The snag is that the cause is unknown, but that the incubation-period is long, perhaps 20 years. Those wishing to make precise the risks of present British dietary practice may have to wait that long before they know.

The effectiveness of present precautions in Britain against the transfer of the BSE infective agent from cattle to people is another source of difficulty — made no easier by the latest development reported on page 280 in this issue. The rules are that animals suffering from the disease must be slaughtered and their carcasses destroyed; the brains and spinal cords of other cattle sent for slaughter must be removed by slaughterhouse dissection, and not committed to the human food chain. A certain cynicism inevitably prevails, perhaps fuelled by the observation, earlier this year, that the rate of diagnosis doubled from one week to the next when the British agriculture ministry increased the rate of compensation for slaughtered animals from 50 to 100 per cent. That does not prove that all farmers are venal, but merely that some are capable of self-delusion.

The danger, in Britain, is that the scare will get out of hand. The public worry about BSE is similar to that about the commercial exploitation of nuclear power. The underlying consideration is the welfare (which, loosely, means good luck in the avoidance of death) of individuals in the population at risk. In each case, putatively damaging agents (prions or radioactive substances) are transferred sequentially from one repository to another (from sheep to cattle to people — possibly — or from reactors to the atmosphere to, say, the lungs). The difference is that, by now, the problems of nuclear power are reasonably well understood. The threat from BSE is, by comparison, almost an unknown quantity. What, in these circumstances, should an agriculture minister do and say?

Gummer, the hapless incumbent of the ministry, is not short of advice on what to do. The most radical suggestion, from Professor Richard Lacey of the University of Leeds, is that all animals from herds in which the disease



has been found should be slaughtered. That is based on the assumptions either that cattle can catch the disease from those with which they are in contact or that those in the same herds have been fed the same foodstuff. That would entail the destruction of some six million animals. That draconian recipe is politically unacceptable. It is probably also unnecessary. More sensible options for action stem from the supposition that maternal transmission happens in cattle as it does in sheep. The British Veterinary Association was advocating last week that there should be a ban on breeding from the progeny of infected animals, while the British Labour Party went one step further and called for the slaughter of all offspring of BSE cattle. In each case, action is impeded by the length of the incubation period of the disease, and by the lack (in Britain) of a system for tracking the pedigree of cattle. If Gummer wishes to be seen to be acting decisively, he could do worse than take himself to Ireland, which has had a workable system of cattle identity for many years. That, of course, might be demeaning for a British mainland politician.

The economic cost of solutions involving the slaughter of 'at risk' animals is a powerful disincentive. At the current rate of compensation (up to £656 if BSE is confirmed, up to £820 if it is not), Lacey's plan would cost the government several thousand million pounds. Even the Labour party's more modest proposal could run up a bill matching the £6 million paid out to farmers so far. But there is a danger that whatever the British government spends, the economic costs to British farmers stemming from the reduced demand for beef will be greater.

What Gummer should say, as distinct from do, is another matter. Even the agriculture minister's fiercest critics seem glad that it is he, not they who is responsible. But there are a few elementary rules that a man in his position may find useful. First, "don'ts". Never say that there is no danger (risk). Instead, say that there is always a danger (risk), and that the problem is to calculate what it is. And never say that the risk is negligible unless you are sure that your listeners share your own philosophy of life.

The "dos" are philosophically more difficult, but for that reason more important. Gummer should be obliged to tell it like it is. Most probably, BSE would not have arisen if farmers had not been helped to feed processed sheep offal to cattle in the early 1980s. But they were, and the British then paid less for beef than might otherwise have been necessary. Now, with the boot on the other foot, the British will not eat beef for fear that it will kill them, and the price has fallen even further.

Gummer should be saying that predicting such an outturn would have required great feats of imagination. He should also be emphasizing the literally unknown steps in the causation of BSE. An origin in scrapie? The link with contaminated sheep feed is circumstantial only but there are ways in which BSE arising *sui generis* could be worse. Maternal transmission? If scrapie is a guide, maternal transmission is more likely but is similarly unproven and difficult to prove. Adventitious infection?

Almost inconceivable, given what little is known of prions, but not entirely out of court. The transfer of BSE to people as CJD? There is not even circumstantial evidence, only hypothesis.

It would not be entirely imprudent for Gummer also to juggle with a few numbers. So far, roughly one in a thousand British cattle have been affected. If the origin is scrapie, and if control measures have been properly applied, a similar number of cattle may be affected in the next five years, after which the disease should disappear. If the transfer of the BSE agent to people from cattle has been contained by the stricter regulation of offal, it is hard to think that people will have been more at risk of acquiring the agent from contaminated beef than cattle have been from contaminated feed. And that risk should now have been eliminated. So the risk that a person eating beef in Britain in the past five years would contract CJD would be between zero (if BSE cannot cause CJD) and, at the most, one in a million. That would make 50 cases of CJD spread over the next 30 years or so, compared with the present case load of about 30 cases a year. Such numbers, only illustrative, might help to take some of the anxiety out of what will otherwise be a bruising debate for all concerned. □

Summit's agenda

Next week's meeting of US and Soviet presidents will be good, but could have been better.

THE Washington summit next week will go as well as can be expected. The two superpowers (if that remains the right word) will have made substantial progress towards a treaty on the use of strategic weapons. After fine-tuning on matters such as cruise missiles it will be a valuable piece of paper, when signed later in the year. It is not so much that it will help to reduce the risks of nuclear war, but that it has been made possible by the natural reduction of these risks by other causes. That is worth remembering when both sides set out to tell what difference its treaty may make to their domestic money problems.

The disappointment of next week's summit will be the two sides' failure to reach an understanding on conventional forces in Europe. This has been on the cards since Mr Mikhail Gorbachev's declaration two years ago that Soviet armed forces would be reduced by 500,000. It has been brushed aside by the rapid pace of political change in Europe — notably by the *de facto* collapse of the Warsaw Pact. The Soviet Union is now dragging its feet for one plain reason — it does not wish to get rid of conventional forces while the military status of Central Europe is undeclared — and one dark reason, mounting domestic opposition to Gorbachev's policies. The summit next week could most usefully lay the basis of the long-term security of Central European states shaken loose from the Warsaw Pact. That Europe's most urgent problem is also its most difficult is no excuse for neglecting it. □

MITI maps out the sixth generation computer

- Fuzzy networks a key component
- Project may be international

Tokyo

JAPAN'S Ministry of International Trade and Industry (MITI) has taken its first tentative step towards mapping out plans for a follow-up to its fifth-generation computer project which ends next year. Details of the new project remain obscure (and in fact are undecided) but it is clear that MITI has set its sights on developing a new generation of massively parallel computers with 'flexible' brain-like information processing systems, including both neuro- and optical computers.

The goals of the new project, which is expected to start in 1992, are described in a report released last month by MITI's Research Committee on New Information Processing Technology, a committee of more than 100 scientists drawn from industry, universities, and MITI's Electro-technical Laboratory (ETL) in Tsukuba.

Unlike the fifth-generation computer project, which concentrated on the development of a highly parallel computer based on existing predicate logic, a key aim of the new project will be to lay down new theoretical foundations for parallel computing. Shunichi Amari of Tokyo University, who developed some of the mathematical theory behind neural network models in the 1960s and 1970s, is an adviser for the project and a strong advocate of such theoretical work.

The theoretical part of the project will set out to establish 'soft (flexible) information processing systems' capable of handling ambiguous or incomplete information in the form of patterns or symbols as well as numbers, with the capacity of learning and self-organization, and that can deal with "approximately correct problem solving" and the "integration of mass information". The research will include and expand on the theory of fuzzy computing and neural network processing, according to Taizo Nishikawa, deputy director of MITI's industrial electronics division.

On the technological side, the project will develop massively parallel computers by looking at various types of new devices, computer architecture and software. The committee has not set its sights on a specific type of parallel computer.

The report includes diagrams of an optical digital computer and a picture of a small piece of hardware that contains one thousand million 'neurons' and a hundred million million transistors on 8-inch wafers stacked 50 centimetres high on a board only one metre square (the commit-

tee sees wafer-scale integration as one possible route to the development of massively parallel computers). But with a power consumption of 160 kW, cooling such a pile of hardware will be just one of many technical problems to overcome.

Although MITI officials are not prepared to comment on the size of budget for the project, it is widely expected that it will be similar in scale to the fifth-generation computer project which will have consumed about ¥50,000 million (\$325 million) over ten years by the time it ends next year.

ETL researchers who were involved in planning both projects are determined that the new project should not be organized along the same lines as the fifth-generation computer project. Their reason is simple. ETL did not get a penny out of the fifth-generation computer. (Similarly, people in industry are not happy because their best people have been tied up in the project for small financial returns.) ETL belongs to MITI's Agency of Industrial Science and Technology. But the fifth generation computer project comes under the jurisdiction of MITI's Machinery and Information Industries Bureau (to which Nishikawa's industrial electronics division is attached). Such is the internal bureaucracy of MITI that despite early attempts ETL could not receive any fifth-generation computer funds.

One way to surmount infighting might be to make the new project international. Nishikawa in fact says that it is "imperative" that the project be international, and he and some of the committee members have already visited government officials and academics in the United States and Europe to explain the goals of the new project. Nishikawa hopes to include foreign representatives (from both government and industry) who will be based in Tokyo as many of the present committee members would prefer to communicate in Japanese.

Nishikawa expects that the new international planning committee will hold its first meeting in June or July. The main aim of the committee will be to decide how the goals mapped out in last month's report can be achieved. Next year the project will enter the 'feasibility study' stage and although Nishikawa is not prepared to comment on the budget for fiscal year 1991, ETL researchers expect the project to get several hundred million yen (a few million dollars).

David Swinbanks

ENVIRONMENT

Compromise at Bergen conference

London

ALTHOUGH dismissed as vague by environmentalist groups, the ministerial declaration from last week's conference on sustainable development in Bergen, Norway, should help progress towards the more definite international agreements on environmental protection expected to be reached at the global United Nations conference in 1992.

In line with Norwegian proposals, delegates agreed that lack of "full scientific certainty" over threats to the environment should not postpone preventive measures. But, at the insistence of the United States (see *Nature* 345, 193; 17 May 1990), commitments to rigid targets for reduction of carbon dioxide emissions and for financial aid to developing countries to phase out chlorofluorocarbons (CFCs) were excluded.

Ministers from 34 countries (including Western and Eastern Europe, plus the Soviet Union, the United States and Canada) agreed to draw up national strategies for carbon dioxide controls after they receive the report that will soon be completed by the Intergovernmental Panel on Climate Change (IPCC). The report will be considered next week by the scientific working group of IPCC at a meeting near Windsor in the United Kingdom. A draft of the report has already been leaked to the media (see *Nature* 344, 577; 2 April 1990). The Windsor meeting will be followed in the next few weeks by meetings of the working groups on the impact of climate change and policy to combat climate change in Moscow and Geneva.

UK environment minister David Trippier, attacked before the conference over British support for the US position on carbon dioxide emissions, said the United Kingdom will establish its national strategy before the world climate conference in Geneva in November. But he warned that solving the greenhouse problem could cause "pain and anguish" to the British electorate.

At Bergen, the United States was once again attacked for its reluctance to help developing countries meet the costs of using alternatives to CFCs, just as it was at the Geneva conference on CFCs earlier this month (see *Nature* 345, 193; 17 May 1990). The Bergen declaration calls on the Montreal Protocol meeting in London next month to strengthen international action to protect the ozone layer, but inserts "for example" before the key phrase "through additional resources and technology transfer".

Observers hope the compromise wording will avert the threatened Indian boycott of the London meeting. India and many other developing nations have not yet signed the Montreal Protocol.

Peter Aldhous

New fears on transmission

London

THE UK agriculture ministry's Central Veterinary Laboratory in Weybridge, Surrey, is examining a suspected case of bovine spongiform encephalopathy (BSE) in a cow born after the government banned the use of cattle feed containing sheep and cattle offal. If confirmed, the case will embarrass the UK government, which maintains that the ban prevented any further spread of the disease.

The feed ban came into force in July 1988 after scrapie-infected sheep tissue was identified as the probable source of the BSE outbreak. Confirmation of the new case (believed to be a 15-month-old cow from southern England) may mean that the BSE agent can be transmitted from infected cows to their offspring, probably via the placenta, just as scrapie can be 'vertically transmitted' in sheep. But violation of the feed ban by a farmer left with stocks of feed containing sheep offal is another possible cause. If vertical transmission is to blame, then more cases of BSE in young cows should emerge over the coming months. Fears of vertical transmission of BSE lie behind calls for a ban on breeding from the offspring of animals that suffered from BSE.

Francis Anthony, chairman of the British Veterinary Association's farm animals committee, says that he supports a breeding ban, not on the grounds of public health, but to help the British beef industry. If herds are bred that have no connection with the BSE outbreak, he says, then British beef may regain some lost export markets. But a breeding ban may be limited in its effect: accurate pedigrees are not kept for all British

cattle, and many cows may be infected that do not yet show symptoms.

Anthony estimates that vertical transmission of BSE could delay the eradication of the disease from British cattle by 20 years. Even without vertical transmission, most experts expect new BSE cases to arise until the end of the century, because of the disease's long incubation period.

Anthony believes that the imposition of a breeding ban could give Britain a head start in dealing with BSE, as he thinks it likely that outbreaks will soon be confirmed in other countries (see *Nature* 344, 805; 26 April 1990). In the United States and France, BSE cases may be dismissed as rabies, Anthony says. The United States Department of Agriculture (USDA) seems to share his concern. Gary Colgrove of USDA says that a project has just been begun to examine the brains of cattle with unidentified neurological disease. Those suspected of containing lesions caused by BSE will probably be sent to the United Kingdom for further examination.

The current uproar surrounding BSE in the United Kingdom was sparked off by the discovery of spongiform encephalopathy in a domestic cat from Bristol (see *Nature* 345, 194; 17 May 1990), which reawakened fears of transmission to humans. A second suspected case, in a cat from Northern Ireland, is now being examined. Agriculture minister John Gummer, facing growing criticism from the British press, appears before the House of Commons agriculture select committee this week to explain the British Government's action to control the BSE outbreak.

Peter Aldhous

Health worries over use of milk hormone

Washington

COMPLAINTS by small dairy farms that a genetically engineered hormone designed to boost milk production could put them out of business has led two US states temporarily to ban the product.

Last month the governors of Minnesota and Wisconsin signed legislation that places a ban until mid-1991 on the sale or use of bovine somatotropin (BST) in the two states, which, combined, produce 20 per cent of the United States' milk. Wisconsin Governor Tommy Thompson said that the ban would allow "additional farmer and consumer education" without hurting the state economically. Vermont's legislature last week rejected a similar proposal.

Opponents of the hormone claim that it encourages overproduction, which would flood the market and eventually drive out small dairy farmers.

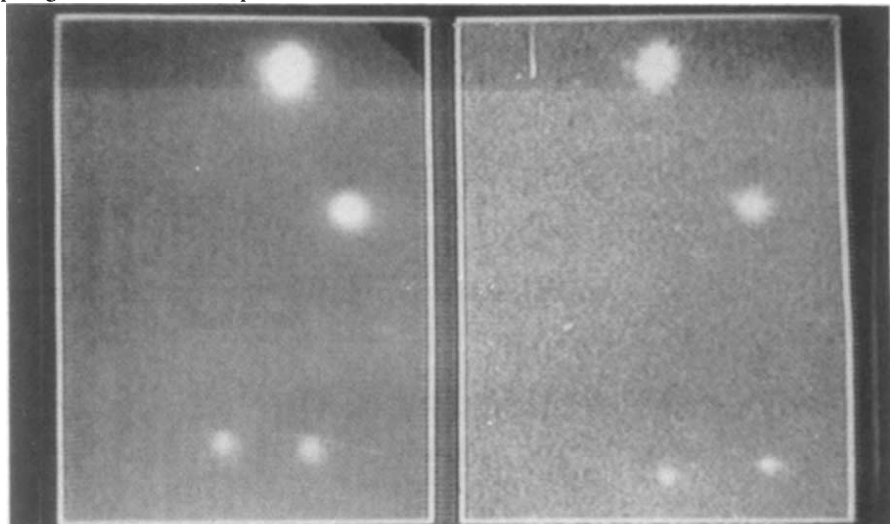
Health worries remain an issue in the debate. Although the US Food and Drug Administration (FDA) says that its studies show that the hormone presents no danger to the public, it has not yet published the data behind that claim.

The US General Accounting Office is investigating charges of favouritism in the FDA's examination of the drug. Jeremy Rifkin, president of the Foundation for Economic Trends, a Washington-based group that opposes BST and some other biotechnology products, points to several studies that raise questions about the hormone's biological impact. Beyond any possible impact on humans, BST appears to shorten the life of the cow. Although milk production can rise 10 to 25 per cent, "cow burnout" can cut reproductive lifespans by a third, he says.

Four biotechnology companies — Monsanto, Eli Lilly, American Cyanamid and Upjohn — have spent more than \$500 million in developing the drug. But BST has been plagued by public opposition and controversy since field trials began in 1985. Last year 2,500 US supermarkets agreed to boycott BST-produced milk (*Nature* 340, 667; 31 August 1989). Until FDA approves BST for general sale and use, which the agency plans to do later early next year, the supermarkets say they will not sell the milk.

Although the BST bans appear to place yet another hurdle in the path of agricultural biotechnology products, analysts caution against extrapolating. "It's one specific product. We don't think it means big trouble when you look at all the other [agricultural] biotech products", says Alan Goldhammer of the Industrial Biotechnology Association.

G. Christopher Anderson



These two photographs of constellation Carina — at right, the first photograph from the Hubble space telescope, and at left, the same star cluster taken by the Las Compañías Observatory in Chile — were released by the US National Aeronautics and Space Administration (NASA) last week. What appears to be an elongated blur at the top of the left photo has been resolved into a double star by Hubble. Although the first Hubble pictures were several times sharper than expected, NASA engineers expect to be able to improve Hubble's resolution considerably when the telescope is fully tuned up later in the year.

Argonne treads carefully

Washington

IN an arrangement carefully crafted to avoid charges of federal meddling in industry, an investment fund set up by the Argonne National Laboratory and the University of Chicago has been used to start a company to market the laboratory's superconductivity technology. Officials say the arrangement marks the first time a US government laboratory has become involved in the creation and management of a private company.

Known as the Illinois Superconductor Corporation (ISC), the start-up company obtains one third of its initial \$1.5 million funding from the ARCH Development Fund, a venture-capital fund started by the laboratory and the University of Chicago, which runs Argonne for the Department of Energy. Additional funds come from the state and a private venture capital firm. The company is the first Argonne spinoff created by ARCH.

Although laboratory officials are careful to emphasize that federal dollars are not going directly to the company, but are filtered through ARCH, the agreement is reminiscent of a deal that lost the chief of another federal agency his job last month. Craig Fields, head of the Defence Advanced Research Projects Agency (DARPA), was removed from his post after deciding that the agency should invest \$4 million in Gazelle Microcircuits Inc., a small semiconductor company.

For the conservative Bush administration, which favours free markets and

resists government intervention in industry, Field's move smacked of an "industrial policy" by which the government would decide to support one technology or company over another.

In the case of Gazelle, the technology was fast gallium arsenide microelectronics, a difficult but promising technique that has had trouble attracting traditional investors. By becoming a partner in Gazelle, DARPA — and the federal government — effectively became a very large and potentially unfair competitor in the industry.

What makes the Argonne start-up different, says ARCH director Steven Lazarus, is that the laboratory is driven solely by a wish to get its technology into the field, rather than the urge to support an ailing part of the industry. Six Argonne superconductor inventions will go into ISC's first product, a sensor to measure the temperature of refrigerants. Because the independently run ARCH — and not the laboratory directly — is the actual partner in the new company, other companies are unlikely to complain that they are forced to compete against the US government, Lazarus says.

Argonne puts \$200,000 a year into ARCH to find outlets for the laboratory's technology. If ISC becomes successful, a share of its profits will return to ARCH, which will then return funds to Argonne's technology transfer programme to encourage further spinoffs.

G. Christopher Anderson

Japanese/US joint venture announced

Washington

CONSOLIDATION within the biotechnology industry continues apace with the announcement this week that US-based Genetics Institute is entering into a joint venture with Yamanouchi Pharmaceutical Company Ltd, one of Japan's largest drug companies. The two-part agreement involves the formation of a separate company for the commercialization and marketing of Genetics Institute's genetically engineered bone morphogenetic proteins (BMPs) in Japan, and the setting up of a partnership to support the development and commercialization of BMPs worldwide.

As a result, Genetics Institute is likely to become the first US biotechnology company to establish a direct marketing presence in Japan. In the carefully crafted agreements, Genetics Institute will retain worldwide manufacturing rights and US marketing rights to its BMP products, and will receive a 50 per cent share in the profits from the sale of its BMPs in Japan.

Gabriel Schmergel, president and chief executive officer of Genetics Institute, expects that by the late 1990s, up to one-third of Genetics Institute's business could be in the bone area, worth in excess of \$2,000 million worldwide. Bone morphogenetic proteins have been shown to induce new cartilage and bone formation. They have potential as a substitute for existing bone graft materials in fracture healing and in the treatment of bone loss associated with periodontal disease and certain cancers.

BMPs are in the pre-clinical testing phase of development and products are not expected on the market for some years.

Genetics Institute has forged close business ties with Japanese companies — Chugai Pharmaceuticals and Suntory as well as its new joint venture with Yamanouchi. "All of our past deals in Japan were licensing deals — everything else will now be channelled through this joint venture, of which we own 50 per cent", says Schmergel. Diane Gershon

Soviet doors opening up

Leningrad

IT has been cogently argued that of the many frustrations to research scientists in the Soviet Union, the worst is the sluggish flow of information from elsewhere: in a fast-moving field, not to know what others have already done, or tried and failed to do, is crippling. While the post, telephone and fax on the whole serve Soviet scientists rather ill, perhaps the most important device for the transmission of information is the international conference. So it is a significant achievement of the Institute of Cytology of the Academy of Sciences of the USSR to have brought to Leningrad last week some two dozen biologists from the United States, Europe and Japan (but mostly from the United States) to speak, along with about a dozen Soviet participants, on the highly topical theme of chromosome transmission and mitosis.

The idea for the conference seems to have arisen during a year-long visit by one of the organizers, Natalay Kouprina (Institute of Cytology, Leningrad) to the laboratory of co-organizer Phil Hieter at Johns Hopkins University in the United States; other members of the organizing committee were Vladimir Larionov and N. Nikolski at the Institute of Cytology, and William Earnshaw at Johns Hopkins.

Scientifically, the conference marked the early stages of a new assault on the macromolecular miracle whereby each time a eukaryotic cell divides, its microtubules assemble into a mitotic spindle that captures the paired and duplicated chromosomes and delivers one set reliably to each daughter cell.

Two traditional lines of attack on the problem may be expected to converge: many participants reported the identification of mutant yeast genes that affect the mitotic machinery so as to increase chromosome loss; many others contributed animal proteins that bind to the chromosomes themselves, or to the spindle microtubules, or both, and are thereby implicated as components of the machinery. The expectation is that some of the animal proteins will turn out to be the products of the genes defined in yeast.

The conference provided a forum for discussion of a central topic in cell biology at an exciting phase of its development when much of what is happening is not yet in print (but see *Nature* 345, 22 & 345, 206; 3 & 17 May 1990). This is the sort of thing Soviet scientists need. But it cannot have been conducive to Soviet participation in informal discussion that, in the retreat to the hard-currency hotel bar each evening, local participants were put in the invidious position of having to accept the hospitality of their guests.

Miranda Robertson

Neurotoxicology in need of a fix

Washington

US research into the neurotoxic effects of hazardous chemicals is fragmented and inadequate, concludes a report released last week by Congress's research body, the Office of Technology Assessment (OTA). As a result, illness stemming from exposure to neurotoxic substances may often go undetected.

Among the options the report offers are substantial increases in support for research and screening of potentially harmful compounds, increased support for neurotoxicology training programmes and a five-year plan to boost the impact of research supported by the National Institutes of Health (NIH).

OTA found the programmes currently carried out under the auspices of NIH, the Environmental Protection Agency and the Alcohol, Drug Abuse, and Mental Health Administration to be of "insufficient size and scope" to address the full problem of environmental neurotoxins. The report stresses that EPA's intramural neurotoxicology programme has been hampered by lack of funding. Despite the need for effective *in vitro* neurotoxicity tests that will minimize the use of animals, EPA recently had a request for extra funds to develop such tests turned down by a government budget office.

The new report emerges against a background of increasing concern over the neurological effects of chemicals. In a series of influential studies, William Langston and colleagues at California Parkinson's Foundation found that MPTP, a chemical created as by-product during the illicit manufacture of synthetic heroin, can induce the symptoms of Parkinson's disease. MPTP's development as a herbicide had previously been halted only by the emergence of more effective chemicals. Langston now believes that the "preponderance of the evidence" indicates that Parkinson's disease is largely caused by non-genetic factors, although others still think there may be a significant hereditary contribution.

Also mentioned in the OTA report is evidence suggesting that a recent increase in the incidence of motor neuron disease, primarily amyotrophic lateral sclerosis (ALS), in the United States is due to environmental agents. But establishing the identities and biochemical effects of the causative agents is likely to prove difficult. The report claims that differences in scientific judgement on the general role of environmental factors in diseases such as ALS and Alzheimer's disease, as well as the parts played by particular causative agents in the progress of neurological diseases, hamper the coordination of federal research and regulatory programmes.

David Concar

Gore weathers storm

Washington

THE latest in a series of sporadic legal wrangles over the true identity of Gore-Tex — a stretched form of the low-friction plastic Teflon which is widely used in outdoor clothing as well as such things as space suits and artificial arteries — has ended with the US company that discovered the material coming unstuck over a vaguely worded patent and evidence of an earlier discovery in Japan.

The company, Maryland-based Gore Associates Ltd, has long weathered legal storms over Gore-Tex, but last week showed signs of losing its grip on the material when a US federal court in Phoenix rejected the company's claims of patent infringement by Impra, a medical prosthetics company in Tennessee. Judge Charles Hardy, ruled that Gore's patent specifications failed to define a unique material and so offered no protection.

Gore-Tex was discovered serendipitously by Robert Gore in 1969 when in stretching a rod of polytetrafluorethylene (PTFE) — the plastic from which Teflon is made — he inadvertently gave it a sharp tug rather than a slow pull. The resulting material was an expanded form of PTFE that combined the chemical inertness and low-friction properties of Teflon with high tensile strength and permeability.

What began as a small basement business is now an international concern with a reputed turnover of more than \$700 million. Yet Gore's success belies a host of patent problems. It took three failed applications and a delay of ten years before the company succeeded in obtaining a US patent in 1979. And in 1984, the company came close to losing that patent in a suit against a New York company called Garlock Inc. Throughout its patent wrangles, Gore has been dogged by questions about an earlier Japanese patent issued to Sumitomo Electric Industries Ltd in 1967. In 1984, Garlock Inc. tried but failed to convince an appeals judge that the materials specified by the Japanese and Gore patents were indistinguishable, and that the later Gore patent was therefore invalid.

Gerald North, Impra's attorney and a veteran of the 1984 dispute, believes that one reason why Impra has succeeded where Garlock failed was that Impra commissioned a team of distinguished materials scientists — led by Edward Clark of the University of Tennessee who pioneered the early studies of PTFE in the 1950's — to make the Japanese material, following the patent specifications down to the last detail. The result, says North, was a plastic which, when tested according to Gore's own protocol, proved indistinguishable from Gore-Tex.

Gore's attorney, Iain Campbell, however, maintains that the prior Japanese

patent is "hopelessly vague" and that it was only with the "hindsight" afforded by the later Gore patent that Impra's scientists were able to use it as a recipe for Gore-Tex.

A second factor crucial to Impra's victory was the judge's ruling that Gore's patent fails to define which of several possible tests should be used to characterize the material's tensile properties. "Gore had provided in effect no way to determine whether infringement had or had not occurred in a particular case", says North. But Campbell argues that the ruling on tensile strength tests reflects a "wooden interpretation" of the patent and that it would be obvious to an expert which of the tests should be used.

Gore-Tex intends to challenge the ruling in the Courts of Appeal. And according to Campbell, if the ruling is upheld on appeal it could have general implications for the way in which patents for new materials are written in the future.

David Concar

ARIANE

Back up by August

Paris

THE European commercial space rocket, Ariane, should resume launches at the end of July or the beginning of August, according to Frédéric d'Allest, the director of Arianespace. Launches were stopped pending an inquiry after the last launch, flight 36, ended in an explosion a few minutes after lift-off. As the cause of the explosion was a discarded rag blocking cooling circuits, safety checks have been stepped up, but the design of Ariane rockets has not needed to be revised.

Customer confidence in the launcher has evidently not been shaken either — six new contracts have been signed since the explosion. The next launch will carry the French direct-broadcasting television satellite, TDF-2, and DFS-2 Kopernikus, a West German telecommunications satellite.

Peter Coles

FAST-BREEDER REACTORS

Superphénix encore

Paris

THE French fast-breeder reactor, Superphénix has developed a new leak in its liquid sodium cooling system. At fault is a junction in heat-exchange circuitry, and about 10 litres of coolant are said to have been lost during the night of 27 April.

The reactor has been shut down while tests are carried out for other cracks and is expected to restart in June. The reactor, for whose closure neighbouring Swiss residents have been campaigning for several years, developed a serious leak in 1987 and was out of service for two years.

Peter Coles

INDOCHINA

Cambodia's killing fields

Annapolis, Maryland

TEN years after the Khmer Rouge were driven from Phnom Penh, rice research and production in Cambodia remains a victim of the Pol Pot regime, Glenn Denning of the International Rice Research Institute told a Rockefeller Foundation agricultural research conference last week.

Among the scars left by the Khmer Rouge attacks of the early 1970s are burned rice-research stations like the one shown here. As the Khmer Rouge first advanced towards Phnom Penh, Cambodia's rice farmers abandoned their crops to flee to the cities. By 1975, when Phnom Penh fell, rice-growing area in Cambodia was a fifth of the 2.5 million hectares cultivated five years earlier, Denning said.

In Cham Car Daung, near Phnom Penh, the Khmer Rouge took over the Institute of Agriculture, the country's main agricultural research university. It was converted into an ammunition dump by the Pol Pot regime, then blown up on their departure in 1979. Of the 300 graduates from the institute's heyday between 1969 and 1975, only 20 have been accounted for. The rest have either 'perished' or left the country, Denning said.

To convert Cambodia to irrigation, Pol Pot banned traditional deep-water rice and commissioned massive networks of waterways. 'Pol Pot canals' still mark much of the country, but few are functional. Those that are still maintained are used as fish ponds.

During the Khmer Rouge attacks, hungry Cambodians ate their seed rice, losing the ability to plant again. With the



The legacy of the seventies.

lost fields went hundreds of unique rice species. "It appeared that centuries of selection had been lost," Denning said.

Since 1981, shortly after the first relief workers arrived in Phnom Penh, the International Rice Research Institute (IRRI) based in the Philippines has reintroduced 562 traditional species to Cambodia, Denning said. Half of the rice now grown in Cambodia comes from species taken from IRRI seed banks. IRRI currently has three scientists stationed in Phnom Penh, including an anthropologist to study traditional farming methods.

After a decade of collective farming in 'crop solidarity groups', Cambodian farmers were last year given the option of 10–20 year leases on the land they worked and were allowed to market their own crops. Although prices of both rice and fertilizer have subsequently gone up, so has production. Rice-growing land now accounts for 1.8 million hectares. Even so, average yield is only 1.2 tons per hectare — still the lowest in Asia.

G. Christopher Anderson

PASTEUR INSTITUTE

Revival of links with Romania

Paris

AFTER 15 years of imposed isolation, the Instituto Cantacuzino in Bucharest has been able to renew contact with the Pasteur Institute in Paris, the institute upon which it was originally modelled. A package of initiatives worked out by the Pasteur Institute aims to bring researchers up to date in the latest techniques of molecular biology and modern virology and to help put the Institute back on its feet.

Set up in the 1920s, the Instituto Cantacuzino should be playing a central role in efforts to control the spread of hepatitis B and AIDS. But after a decade without access to international science, either through the literature or through travel, researchers have been left far behind.

Training will be offered in Paris to Cantacuzino's best researchers and a 3-week summer school organized in Romania, for about 20–25 researchers.

One of the most urgent needs, says Marie-Hélène Marchand, secretary-general of the Pasteur Institute, is for books and journals. The Pasteur is currently trying to put together complete series of the major research journals for the past decade and to provide subscriptions for this year.

Another initiative, at the request of the Instituto Cantacuzino, is to administer funds it has received from abroad (notably from Belgium). Products and equipment will be ordered and paid for in France and sent to the institute, which lacks even the bare essentials for accounting.

The Cantacuzino Institute has historic links with the Pasteur and its founder, Jean Cantacuzino, was himself a Pastorian. Its present director, Auber Combiescu, grew up in France. Romanian-born researchers at the Pasteur Institute have been instrumental in reopening contacts.

Peter Coles

WORLD HEALTH ORGANIZATION

New man takes over

London

MICHAEL Merson, a US citizen, is the new head of the World Health Organization's global programme on AIDS. He replaces Jonathan Mann, who resigned in March after disagreements over policy with WHO director-general Hiroshi Nakajima (see *Nature* 344, 283, 22 March 1990).

Merson has headed WHO's diarrhoeal disease-control programme since 1984, and also took control of its acute respiratory infections programme in 1987. Addressing the 43rd World Health Assembly in Geneva after his appointment, Merson said that WHO's continued cooperation with other non-governmental organizations (NGOs) is essential in the fight against AIDS. Nakajima is believed to have opposed Mann's moves to increase links between the AIDS programme and other NGOs. But Merson's position on the distribution of drugs in developing countries, understood to be another source of disagreement between Mann and Nakajima, is less clear.

Merson's immediate concern is the AIDS programme's budget, planned for \$100 million in 1991. But only \$70 million has been set aside so far. Merson expressed "great concern" if this shortfall cannot be made up, and warned against "complacency" over the global AIDS pandemic.

Peter Aldhous

IN VITRO FERTILIZATION

Children 'normal'

London

CHILDREN conceived by *in vitro* fertilization (IVF) seem no more likely to suffer congenital abnormalities than the general population, a Medical Research Council working party has concluded. But the infant mortality and still-birth rate in IVF children is twice the UK average because of the high proportion of multiple births.

The survey looked at 1,267 pregnancies in the United Kingdom from 1978 to 1987, conceived by IVF or gamete intrafallopian transfer (where sperm and eggs are deposited in the fallopian tubes), which resulted in 1,581 live and still-births. Its publication in last week's *British Medical Journal* is timely, because the Human Fertilization and Embryology Bill, currently passing through parliament, will set up a Statutory Licensing Authority to regulate IVF.

The present Interim Licensing Authority guidelines recommend that three IVF embryos should be placed into the uterus to ensure a reasonable success rate. But this increases the rate of multiple births. In the opinion of Brian Lieberman, from St Mary's Hospital in Manchester, improved preservation of frozen embryos over the next two or three years should mean that multiple births will be common: if spare embryos are preserved, then there will be less initial pressure to obtain successful implantation.

Peter Aldhous

A friend from Wall Street

Prague

GEORGE SOROS, an expatriate Hungarian who made millions on Wall Street, is to put up \$25 million as seed money for a new English-language university in Prague through his New-York-based 'Open Society Fund'. If delicate negotiations with the interim Czechoslovak government can be completed by the end of this month the university, the first of its kind in Central Europe, may open its doors as early as 1991.

Soros, who is 59, has supported cultural and educational causes in Hungary since 1984 by providing nearly \$3 million a year through the 'Soros Foundation — Hungary', an offshoot of the Open Society Fund. Before the recent ousting of the Communist government, the foundation was known in Budapest as the 'alternative ministry of culture'. The new 'Central European University,' will offer post-graduate courses of up to three years and "support the transition to a market economy", according to physicist Petr Pajas, director of the project in Prague. Although the initial focus will be on the social sciences — a field that was especially deformed by Communist governments according to Pajas — the university may eventually offer courses in natural sciences such as ecology and 'evolution theory' which were also neglected.

The lone remaining obstacle to the project is finding a building in Prague suitable for intensive use as a teaching and research centre. Pajas hopes that the negotiations with the government will proceed smoothly so that the tasks of finding faculty and supplies can begin next month.

The \$25 million from Soros will be used to attract faculty members, to set up special short courses in the 1990–91 academic year and to lay the groundwork for the full-scale start of the university in the fall of 1991. Fifty students will be taken in at first but Pajas expects numbers to rise to 300. Intensive training in English will be offered at least in the first two years.

Although the building will probably come from the government, Pajas stresses that the goal is "to keep the university as independent of the state as possible". The university will seek private donations for computers, books and other supplies.

An additional site has been chosen at Bratislava, the capital of Slovakia, just east of Vienna. Although the main campus will be in Prague, the full university may eventually feature satellite campuses in Budapest — where Soros has already set up an international management school with Western professors — Vienna and Graz.

Steven Dickman

Self-help plan

London

THE communist legacy of poor training and research management in the environmental sciences compounds Czechoslovakia's pollution problems, a Czechoslovak delegation revealed last week. The delegation, led by Zdenek Kukal, research director of the national geological survey, was in London to formulate proposals for support from Britain's 'Know How Fund' for Czechoslovakia. "This is about helping the Czechs to help themselves", says Professor Mike Boulter, from the Polytechnic of East London, who organized the meeting.

Karel Urban, from the Ministry of Culture and Nature Protection, explained that ecology had been viewed as an 'anti-socialist' subject. Data on Czechoslovakia's environmental crisis have now been released, and the powerful green lobby is demanding urgent action. But the newly formed environment ministry is financially stretched, lacking even a central headquarters building.

Czechoslovakia has enough graduates in the basic biological sciences, Urban said, but there is a shortage of environmental scientists. Other delegates berated the traditional bias against applied environmental research, and argued that a new system of research management is needed. "We are amateurs in managing research", Kukal said, and suggested that a group of Czechoslovaks should visit Britain to learn from the British system.

Viktor Hoschl, a self-labelled "child of the revolution" and geophysicist with Geoindustria (a state-owned company), said that exchange of information with the West is needed to help Czechoslovakia deal with its accumulating nuclear waste. Under communist control, waste was sent to the Soviet Union.

Peter Aldhous

New laws for old

Prague

Two pieces of legislation of vital interest to the scientific community are among the 63 new laws passed by the Czechoslovak parliament since its reconstruction at the beginning of the year. But there is general anxiety about one of these bills, the law on the universities and higher education, rushed through before the end of the present session.

Essentially, the new law gives more autonomy to university rectors and deans, allows students to participate in university government and opens teaching posts to competition. The fear is that the slapdash way in which the law was drafted may conceal snags that defeat aspirations for the future.

By contrast, the amendment to the law on the Czechoslovak Academy of Sciences, was passed without excitement. The amendments give the elected representatives of academy institutes rights equivalent to those of members of the academy. The new law also abolishes the requirement that academy scientists should compete for their jobs every five years, which had been used for blackmail by the previous regime. However, membership may now be terminated on the grounds of unethical conduct or in cases of members appointed for "non-scientific" reasons.

The legislation also legitimizes the way in which the academy has operated since last December, since when a panel of seven elected scientists has shadowed the decisions of the formal praesidium, exercising a veto on matters with which it disagreed. Now it should be possible for those in charge to turn from criticism of the old order to finding a more logical and less discriminatory way of conducting research.

Veronika Maxova

Lesotho academic excluded

Cape Town

THE decision to refuse a South African residence permit to Professor Njabulo Ndebele, a citizen of Lesotho who was chosen by the University of the Western Cape (UWC) to be its second vice-rector (see *Nature* 344, 581; 12 April 1990) was made at cabinet level, according to a statement released by the Minister of Home Affairs and National Education, Gene Louw, after a meeting with UWC rector Jakes Gerwal and vice-rector Jaap Durand on 17 April.

Louw said that Ndebele, an exiled South African, was placed on the government's list of persons to whom visas should not be granted several years ago, and that it was "not customary" for the government to provide reasons why individuals' names appeared on the list.

Louw claimed that government "was accommodating within all reasonable limits" about requests to grant visas to listed people, but "could not be expected to simply submit to all requests from parties who do not possess all the facts". UWC's council has asked for a meeting with President de Klerk to discuss the case.

Although Louw was dealing with this issue in his capacity as Minister for Home Affairs, his action bodes ill for Louw's future relations with the universities as Minister for National Education, his second cabinet portfolio. Louw took over the post last month from Dr Gerrit Viljoen, who is concentrating on his responsibilities as Minister for Constitutional Development and chief negotiator with the African National Congress.

Michael Cherry

JAPAN'S ANTI-NUCLEAR MOVEMENT

Yen melts down opposition

Rokkasho

THE opening of Japan's first commercial facility for the enrichment and reprocessing of nuclear fuel and storage of nuclear waste is destined to go ahead despite local and nationwide opposition.

Last week, local government officials, representatives of the facility and fishermen opposed to the nuclear complex in Rokkasho village on the northern tip of the main island of Japan made it clear to visiting members of the foreign press that the huge amounts of money brought in by the project have effectively quashed almost all local opposition.

The Rokkasho facility, which will cost well over ¥1 million million (\$6,500 million) when completed in the late 1990s, is central to Japanese government plans to gain greater independence from foreign suppliers of nuclear fuel. At present nearly all uranium used in Japan's nuclear power plants is of US origin and Japan's spent fuel has to be sent to the United Kingdom and France for reprocessing.

The first facility at Rokkasho, scheduled to open next year, will be a ¥180,000 million plant for the enrichment of uranium. The plant, to be run by Japan Nuclear Fuel Industries (JNFI), a consortium of ten electric power companies, is already 60 per cent complete, according to JNFI officials, and when it reaches full capacity in 2000 the plant will be capable of producing about 20 per cent of Japan's enriched uranium needs.

A low-level waste storage facility also owned by JNFI is expected to open in 1992 at a cost of over ¥100,000 million. With enough space for three million 200-litre drums of waste the facility will be capable of absorbing all low-level waste from Japan's power stations well into the next century. The drums will be buried 12 metres underground, encased in concrete, and monitored for leakage for the next 300 years.

Facilities for reprocessing spent fuel will be built by another consortium of over 100 companies, Japan Nuclear Fuel Service (JNFS), at a cost of ¥840,000 million. JNFS officials hope to start accepting spent fuel from Japanese power stations in 1994 and to begin reprocessing it into uranium and plutonium in 1997 after the fuel has 'cooled down'.

At full capacity, the plant will be able to reprocess the spent fuel from about 30 of the 37 nuclear power reactors currently in operation in Japan. The remaining spent fuel will continue to be sent to the United Kingdom and France for reprocessing.

Japanese high-level waste from the reprocessing plants in Europe will be stored 'temporarily' at Rokkasho for 20 to 30 years from 1993 onwards, according to JNFS officials.

Since the Chernobyl accident in 1986, the Rokkasho facility has become the focal point of a growing nationwide anti-nuclear power movement. And in December the incumbent village chief in Rokkasho, a strong supporter of the facility, was ousted in the village election by Hiroshi Tsuchida who campaigned for a freeze until the safety of the facility is confirmed.

Last week, Tsuchida, a former supporter of the project, said that he is personally convinced of the plant's safety and once he feels that about 75 per cent of the approximately 12,000 villagers are also convinced he will allow commercial operations to begin (the village chief must sign a safety agreement with JNFS and JNFI before commercial operations can start). Tsuchida estimates that only a few hundred locals are vehemently opposed to the plants and he and fellow village assemblyman Saburo Hashimoto, a supporter of the project, say that JNFS and JNFI should carry out more public relations campaigns to win over the large number of people taking a 'neutral' stance. In line with such calls, the two companies plan to open a new ¥3,000 million (\$2 million) public relations centre at Rokkasho next year.

Members of an association of 30 fishermen in the nearby village of Tomari who are opposed to the facility claim that over 50 per cent of Rokkasho villagers are against the project and that many people were deceived by Tsuchida's call for a freeze. But the association says that people cannot openly express opposition because they would be turned away from part-time jobs at farms and the construction site for the facility by village officials who are promoting the project. During the construction peak in the next few years, thousands of locals will be temporarily employed to construct the plants.

And in the Rokkasho area, where men often have to go to Tokyo to find employment as part-time construction labourers, the incentive to refrain from criticizing the project is strong, the Tomari fishermen say. But construction jobs are not the only incentive. Hashimoto says that under a Japanese law designed to promote nuclear power, the Ministry of International Trade and Industry (MITI) will provide Rokkasho with ¥18,000 million (\$120 million) in subsidies over a ten-year period. This amounts to about \$10,000 for every man, woman and child in the village.

The money, which is over four times the village's annual budget, will be used for 96 new village projects including a fish farm, facilities for manufacture of fertilizers, roads, radar facilities for fishermen, a large refrigerated warehouse for storing vegetables, a sports amenity centre and a ski resort. Another ¥18,000 million will

ARCHAEOLOGY

US cracks down on Peru loot

Washington

AFTER a three-year investigation into the desecration and looting of the richest tomb ever excavated in the Western Hemisphere, the US Customs Service has announced a ban on the importation of all archaeological artefacts from the Sipán region of Peru. The US Customs Service said it would seize any artefact that was not accompanied by documentation from the Peruvian government certifying that it had left the country legally. "We are appalled by the desecration of the Sipán Region", said Customs Deputy Commissioner Michael Lane, "but we will end this crime." The discovery in



1987 of the graves of the 'Lords of Sipán', as the site is known, led to the 'Peruvian Gold Rush', a frenzy of looting that robbed the tombs of thousands of priceless pre-Columbian objects. Many of the items eventually made their way to the United States, where they were bought by eager collectors.

In 1988, armed US customs agents searching for illegal Peruvian artefacts raided a group of homes and businesses in Southern California. Among the collectors caught in the raid was physics Nobel laureate Murray Gell-Mann, who had invested tens of thousands of dollars in Sipán art and relics. Gell-Mann then worked with the investigators, and returned several dozen artefacts to Peru.

G. Christopher Anderson

be given to several villages surrounding Rokkasho, Hashimoto says.

The fishermen's association has filed suit in the Tokyo district court against former prime minister Noboru Takeshita for approving construction of the uranium enrichment facility. The land at Rokkasho was bought from local people by the government-backed Mustu Ogawara Corporation during the 1960s, on the understanding that it would be used for industrial development. But the fishermen do not expect success. They are pinning their last hopes on support from a nationwide network of anti-nuclear activists and on the success of an anti-nuclear candidate in the elections for governor of Aomori Prefecture to be held next February.

David Swinbanks

Charting the decline

SIR—The letter from J.F. Lamb¹ has prompted us to examine some data extracted from a recently published study². Table 1 shows country-by-country mean citation rates of papers published in *Nature* during 1981–85 as related to the mean citation rate of all papers published in *Nature* in the same period. All countries publishing more than ten papers are shown.

TABLE 1 Relative Citation Rate (RCR) of papers in *Nature*

Country	RCR	Country	RCR
Sweden	1.92	Poland	0.70
Switzerland	1.78	Italy	0.67
Japan	1.46	Austria	0.67
USA	1.18	Norway	0.64
West Germany	1.17	Canada	0.62
Denmark	1.07	USSR	0.52
Israel	1.05	Australia	0.46
The Netherlands	1.01	Spain	0.46
France	0.92	Brazil	0.41
Belgium	0.79	New Zealand	0.41
United Kingdom	0.77	Ireland	0.35
Hungary	0.75	South Africa	0.23
Finland	0.70	India	0.14

The relative citation rate (RCR) has a value greater than 1.00 when the papers of the country in question are cited above the journal's average and conversely. *Nature* published 8,043 papers in 1981–85 with an average citation rate of 16.63.

As can be seen, Sweden heads the ranking, followed by Switzerland, Japan, the United States and West Germany. The United Kingdom is ranked eleventh. This could be considered to be in line with the reports of Lamb and others on the decline of British science^{3,4} although other studies did not find any statistically significant decline in the indicators of British science during the 1980s^{5–8}. For that reason, we consider the whole problem of the decline of British science is still open to discussion — leaving aside questions such as what is meant by decline⁹ and whether whatever it is correlated with the decline in funding, for example.

Accordingly, we have repeated the same exercise with another 'high quality'

periodical, the *Biochemical Journal*.

Table 2 presents the results, covering 4,475 papers published in 1981–85, with an average citation rate per paper of 6.90. As will be seen, the United Kingdom is ranked in first position.

Our sole purpose here is to suggest that a counting such as that of Lamb is not by itself a sufficient measure of decline.

Bibliometric methods are not the most serviceable tools unequivocally to prove such a decline during the 1980s. A more realistic approach would show² that British science, like science in other countries, is in decline in some fields and on an upward move in others.

Moreover, in any population of papers, the distribution according to quality (citation rate is but a proxy) is highly skewed. Our investigations² indicate that probably the most important factor in improving scientific performance of a country is finding a way to raise the quality and not quantity of the publications by influencing the skew distribution to have a low quality 'tail' as short as possible.

T. BRAUN

*Institute of Inorganic and Analytical Chemistry,
L. Eötvös University, POB 123,
1443 Budapest, Hungary*

A. SCHUBERT

*Information Science and Scientometrics Research Unit (ISSRU),
Library of the Hungarian Academy of Sciences,
POB 7, 1361 Budapest, Hungary*

1. Lamb, J.F. *Nature* **343**, 404 (1990).
2. Schubert, A., Glänzel & W. Braun, T. *Scientometrics* **16**, 3 (1989).
3. Irvine, J., Martin, B., Peacock, T. & Turner, R. *Nature* **316**, 587 (1985).
4. Martin, B.R., Irvine, J., Narin, F. & Sterritt, C. *Nature* **330**, 123 (1987).
5. Leydesdorff, L. *Science and Public Policy* **15**, 270 (1988).
6. Nederhof, J. *Scientometrics* **14**, 475 (1988).
7. Braun, T., Glänzel, W. & Schubert, A. *Scientometrics* **15**, 165 (1989).
8. Braun, T., Glänzel, W. & Schubert, A. *The Evaluation of Scientific Research*, Ciba Foundation Conference, p. 32 (Wiley, Chichester, 1989).
9. Nye, M.J. *ISIS*, **75**, 697 (1984).
10. Babbage, Ch. *Nature* **340**, 499 (1989).

(2) Clearly the term 'economic growth' has a ceremonial value, while 'no-growth lobby' has a pejorative ring. Yet even radical 'greens' are in favour of 'economic growth': the issue is what kind of growth. The article begs several questions. Just what, for example, do the national income accounts at present measure?

(3) While the article seems relatively sanguine about the future of European civilization, there is a remarkable inconsistency inherent in the assertion that Africa, "even before AIDS, seemed destined to be the first continent to be lost to human habitation through wayward mismanagement". Even assuming, out of ethnocentric arrogance if nothing else, that Africans bring the house down on their own heads, an ecological viewpoint forces recognition that our own self-interest is at stake. Africa is part of the global ecosystem.

(4) No responsible scientist claims to be able to predict the future (unless all other conditions are equal), yet the essay skates perilously close to prophecy on more than one front. Chaos theory, if it means anything *vis-à-vis* the environmental policy debate, implies that small and seemingly insignificant variables can have major unpredicted effects. Clearly, some kinds of economic growth could have unpredicted and disastrous ecological effects. Furthermore, even assuming adequate funding for basic science research, there is no guarantee that the knowledge necessary to sustain civilization as we know it will be forthcoming.

(5) Finally, the article ignores a host of so-called science–society issues that such luminaries as Albert Einstein and Erwin Schrödinger have raised.

Rather than being the last word, we hope that the leading article serves the function of opening up and extending the conversation. Perhaps the wide range of articles dealing with various facets of ecology, environment and human impacts that have graced the pages of *Nature* are the beginning of a vital process of public deliberation informed by scientific inquiry.

MAX OELSCHLAEGER

*Department of Philosophy,
POB 13526,
University of North Texas,
Denton, Texas 76203, USA*

Missing woman

SIR—Sachi Sri Kantha (*Nature* **344**, 582; 1990) concludes that 21 women so far have been awarded the Nobel prize. However, at least one woman was missing. In 1909, the Swedish novelist Selma Lagerlöf was the first woman to receive the literature prize.

ÅKE STRID

*CSIRO Division of Plant Industry,
GPO Box 1600, Canberra ACT 2601,
Australia*

Growing dissent

SIR—The debate between advocates of sustainable economics and pro-growth economics is more complicated than implied by the recent leading article, "Doctrinal fallacies of stewardship" (*Nature* **344**, 179; 1990).

(1) The assumption that funding for scientific research and development entails economic growth cannot be cogently defended. While we agree that scientific enquiry is at present underfunded, both in Britain and elsewhere, the explanation cannot be attributed to any single factor.

TABLE 2 Relative Citation Rate (RCR) of papers in the *Biochemical Journal*

Country	RCR	Country	RCR
United Kingdom	1.22	Norway	0.80
Sweden	1.21	Canada	0.79
Switzerland	1.16	France	0.79
Belgium	1.12	Spain	0.79
Denmark	0.92	Austria	0.79
The Netherlands	0.92	Hungary	0.74
South Africa	0.90	Ireland	0.67
West Germany	0.89	Finland	0.66
USA	0.89	Poland	0.64
Australia	0.87	Argentina	0.63
Israel	0.82	USSR	0.49
Japan	0.81	India	0.41
New Zealand	0.80	Chile	0.25

What price the global greenhouse?

Is the global greenhouse potentially a calamity to be avoided at all costs? Or an unavoidable disaster of incalculable cost? One view is that the best greenhouse is that which, in the broadest sense, costs least.

Munich

THERE are two distinct and extreme views of the consequences of the general increase of temperature likely to be occasioned by the accumulation of carbon dioxide and other greenhouse gases in the Earth's atmosphere. One is that even the slightest departure from the present fortuitous pattern of climate will be tantamount to a calamity and had better be avoided. The other is that modern technology is now so sophisticated that even the most severe departures from the present state of affairs can be accommodated.

Both positions are, of course, mistaken. Climatic changes indistinguishable from interannual noise would be of no importance, while changes much more rapid than the timescale on which the world's stock of social capital is renewed — say, a century or so — would clearly be a great economic if not absolute embarrassment. What can be said of more modest and more tenable intermediate positions?

In this connection, one should acknowledge that there is some force in the extreme environmentalists' position that even modest change means calamity. Most calculations of the greenhouse effect yield estimates of the increased temperature caused by a doubling of atmospheric carbon dioxide, which on present trends is expected to happen about the year 2035. The modellers seem to agree that the extra temperature then or some decades later (if the oceans substantially delay events) could be as much as 4° C. Most arguments about the adequacy of the models centre on assertions that ameliorating feedback mechanisms may have been overlooked in their construction. But if nothing else is at stake, the result must simply be that the full rigours of the predicted greenhouse are delayed. That is another way of saying "In the long run, we are all dead".

It is nevertheless worth asking what will happen if the projected change of climate does occur in the next few decades. One possibility is that governments might not act energetically enough, or in time. But there is also a colder calculation to be made: on the face of things, avoidance may be the prudent course, but to what degree?

In some broad sense, this is an economic question. The cost of not avoiding rapid climatic change is that of adapting to the conditions that climatic change would bring. The cost of avoiding the greenhouse absolutely would be that of maintaining

constant the atmospheric concentration of carbon dioxide and other greenhouse gases. It is unlikely that the most economical course to follow would be that of absolute avoidance. Nor is the cost of doing nothing likely to be the smallest. By these criteria, and in an ideal world in which all decisions are taken rationally, the best strategy would almost certainly be a calculated blend of avoidance and adaptation.

The obvious trouble is that such questions cannot be made even partially precise in present circumstances. It is not merely that the calculations have not been carried out, but that the basis on which they might be attempted are essentially undefined. What, for example, is the probable effect of climatic change of the kind predicted on the economic basis of city life? Increased temperatures might for example bring reduced heating costs but increased costs for air conditioning. Where would the balance lie, and would the trade-off be such that net energy consumption and thus ultimate greenhouse-gas production would be increased or decreased? Could the radical redesign of buildings change the balance and ameliorate it and, if so, to what degree? What would be the consequences of markedly higher electricity costs or restrictions on the use of internal combustion engines?

The economic basis of industrial activity is even less thoroughly explored, although some rules of thumb can be inferred from past experience. To the extent that restrictions on the release of greenhouse gases, and carbon dioxide in particular, will appear in the economy as increased energy prices, attempts to regulate the greenhouse gases will have effects not very different from the rapid increases of the price of oil in 1973 and 1979, although the opportunities for substitution (by coal, for example) will be restricted. That suggests that energy-intensive metals (such as steel and aluminium) will be at a disadvantage, while refined products of all kinds will become more expensive relative to services provided by people. But to what degree? And to what extent can novel technology provide alternatives? These questions are imponderable, but nevertheless important.

Unsurprisingly, the one field of economic activity to which some attention has been paid is that of agriculture. The explanation is not difficult to fathom. The effects of climatic change on the growth of

crops is, after all, direct; increased temperature and reduced (or increased) rainfall has a direct effect on the way plants grow. So, for that matter, has the increased concentration of carbon dioxide in the atmosphere, which generally enables photosynthesis to be more rapid, but to an extent that varies from one crop to another.

One of the most elaborate studies so far of the effects of climatic change on agriculture is that reported in last week's issue by R. H. Adams *et al.* (*Nature* **345**, 219; 1990), which vividly illustrates the complexity of even the simplest studies aimed at throwing light on the way in which sub-systems of the modern economy are likely to respond to climatic change. For one thing, the study is concerned only with US agriculture, which is the best understood of any, and with only three crops. For another, while dependent for its calculations on model projections of climatic changes caused by greenhouse gases, it must use a mere handful of experimental data on the increase of the rate of photosynthesis in different crop plants caused by increased carbon dioxide.

The conclusions are also complex, and in some ways unexpected. US agriculture could survive climatic change of the kind that may be caused by a doubling of carbon dioxide in the atmosphere, at least if full credit is taken for the stimulation of photosynthesis by carbon dioxide. There would even be agricultural produce left over for the strategically important export trade. But at a cost. Irrigation needs, already met only at high cost, would multiply. Food consumers would have to pay much more for food relative to other consumer prices. Just how farmers would fare is another matter: the two climate models used in the study suggest that the higher prices hurting consumers would (not surprisingly) benefit farmers, giving them a little extra income.

The authors of the study now say that they are unhappy that the last of these issues has been singled out for special attention in the general press, but is that necessary or even wise? The value of a study such as this is not to predict the future, but rather to provide an understanding of a complex economic field that is rich enough to allow for rational planning, preferably before the reality of the greenhouse is palpable. Does it matter that, among the majority of losers, some may be better off?

John Maddox

A finger on the missing link

Ronald C. Desrosiers

PREVIOUSLY unrecognized and apparently new, AIDS burst onto the scene in the 1970s and 1980s in the United States, central Africa, Europe and elsewhere. But where did the causative agent, a human lentivirus called HIV-1, come from? Although the report by Huet *et al.* on page 356 of this issue¹ provides no final solution to the mystery, the authors have uncovered the most significant clues to date. Their genetic sequence of a lentivirus isolate from a chimpanzee shows it to be much more closely related to HIV-1 than any previous primate lentivirus isolates, yet it is outside the range of variation of known HIV-1 isolates.

One of two possibilities seems likely for the origin of HIV-1. First, the virus may have always been present in the human population, in one form or another, but may have gone unrecognized due either to an extremely low prevalence or its confinement to isolated populations. Some features of modern society, for example re-use of syringes and hypodermic needles, changes in migration patterns with extensive worldwide travel, and sexually promiscuous activities, could have greatly facilitated its spread in recent history. Second, HIV-1 may have entered the human population by transmission of a virus from another species relatively recently in history. Increasing prevalence and high mortality would then be consequences of infection of the new host.

Related lentiviruses have now been isolated from sheep, goats, horses, cattle, cats, monkeys and humans. The simian immunodeficiency viruses, or SIVs, are nonhuman primate lentiviruses that are the closest known relatives of the HIVs.

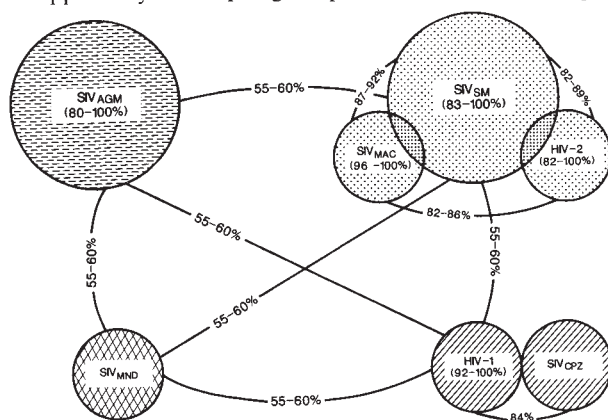
SIVs have previously been isolated from macaques (SIV_{MAC}), sooty mangabeys (SIV_{SM}), African green monkeys (SIV_{AGM}) and mandrills (SIV_{MND}) (refs 2–6). Mangabeys, green monkeys and mandrills are African Old World primates, whereas macaques are Asian Old World primates. These four SIVs fall into three discrete groups based on genetic sequence analysis, with SIV_{MAC} and SIV_{SM} forming a single genetic group^{7–10}. Macaques are apparently

not infected with SIV in their native habitat^{5,11}, so it seems likely that some individual macaques became infected with SIV_{SM} from sooty mangabeys while in captivity at US Regional Primate Research Centers^{9,12}. Genetic evidence as well as geographical considerations provide circumstantial evidence that sooty mangabeys, or some close relative, may also have been a source of cross-species infection of humans for the origin of HIV-2 in West Africa^{9,12}.

HIV-2 is no more different from SIV_{SM} at the sequence level than individual SIV_{SM} isolates are different from each other (refs 9, 12; Y. Li *et al.*, unpublished observations). Furthermore, the natural habitat of sooty mangabeys is the coastal forest regions of West Africa, the same region where HIV-2 is endemic. Thus, this type of molecular detective work has already lent support to the cross-species transmission theory for the origin of HIV-2. But because SIV_{MAC},

SIV_{SM}, SIV_{AGM} and SIV_{MND} are only distant relatives of HIV-1, the source of worldwide human infection with HIV-1 has remained a mystery.

The lentivirus isolate sequenced by Huet *et al.*¹ was derived from a captive chimpanzee in Gabon¹³, and termed SIV_{CPZ}. It is the only lentivirus isolate other than HIV-1 that has a genomic organization that includes the genes *gag*, *pol*, *vif*, *vpr*, *tat*, *rev*, *vpu*, *env* and *nef*. Furthermore, it is much more closely related to HIV-1 at the sequence level than any of the other lentiviruses. For example, amino-acid identities in the *gag* and *pol* gene products are 75 and 84 per



Primate lentivirus taxonomy. Percentages in the lines are amino-acid identities in the *pol* gene product; those in parentheses are the ranges of identities in the product in isolates of each virus.

cent (see figure). By comparison, the other SIVs have only 55–60 per cent amino-acid identity in these products when compared with HIV-1. The other animal lentiviruses have even less (25–30 per cent). Nonetheless, the chimpanzee lentivirus sequence is outside the range of variation of known HIV-1 isolates; even the most divergent isolates of HIV-1 have about 92 per cent positional identity in these gene products. So the chimpanzee lentivirus isolate may be a form of 'missing link' between the African Old World primate SIVs and HIV-1.

These observations do not, however, allow discrimination between the theories for the origin of HIV-1. The data that have been obtained so far are conceivably consistent with both hypotheses. For many virus groupings, related viruses can be found throughout the animal kingdom. Their relatedness to the human virus generally reflects the evolutionary distance of the host species. For example, viruses related to herpes simplex virus (HSV) of humans can be found all through the animal kingdom and their relatedness to HSV in general reflects the evolutionary distance of the host species^{14,15}. Apparently, an ancestor alpha herpes virus was present very early in evolutionary history and the viruses diverged in concert with their host species. The gradient of sequence relatedness from the ungulate lentiviruses, through SIV_{AGM}, SIV_{SM} and

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Snapshooting surface radicals

Martyn C. R. Symons

SIV_{MND}, to SIV_{CPZ} and HIV-1, is consistent with such coevolution. However, if SIV_{CPZ} is part of a large, diverse group similar to SIV_{AGM} (ref. 16) and SIV_{SM} (Y. Li *et al.* unpublished observations), then HIV-1 could also have been derived by cross-species transmission from an individual member of the SIV_{CPZ} group.

Many questions still surround the SIV_{CPZ} isolate. Fortunately, they are amenable to experimental investigation and their resolution will allow us to decide between the competing theories for the origin of HIV-1. Is the virus really a natural infectious agent of chimpanzees? Hundreds of captive chimpanzees have been examined previously and none were seropositive except those intentionally inoculated with HIV-1 or with human materials¹⁷. The chimpanzee from which SIV_{CPZ} was isolated was one of only two seropositives among the 83 tested in the captive Gabon colony¹³. In contrast, 20–50 per cent of African green monkeys in their native habitat are naturally infected with their own SIV. Is there any chance that this chimpanzee became infected from a human source? Chimpanzees occasionally have been observed to eat monkeys in the wild — is it possible that this chimpanzee became infected from some other non-human primate species?

If the virus is indeed a natural infectious agent of chimpanzees, it will be important to determine the prevalence and demographics of infection and to document the extent of genetic diversity among an array of isolates. The status of other great apes will also need to be investigated. If SIV_{CPZ} turns out not to be a natural infectious agent of chimpanzees, it will be important to find the source of infection, particularly if it is some other nonhuman primate. What is clear is that the proliferation of SIV isolations has created a fascinating molecular-genetic game to unravel the origins of the HIVs. □

Ronald C. Desrosiers is at the New England Regional Primate Research Center, Harvard Medical School, One Pine Hill Drive, Southborough, Massachusetts 01772, USA.

- Huet, T. *et al.* *Nature* **345**, 356–359 (1990).
- Daniel, M. D. *et al.* *Science* **228**, 1201–1204 (1985).
- Murphy-Corb, M. *et al.* *Nature* **321**, 435–437 (1986).
- Fultz, P. N. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 5286–5290 (1986).
- Ohta, Y. *et al.* *Int. J. Cancer* **41**, 115–122 (1987).
- Tsujimoto, H. *et al.* *J. Virol.* **62**, 4044–4050 (1988).
- Chakrabarti, L. *et al.* *Nature* **328**, 543–547 (1987).
- Fukushima, M. *et al.* *Nature* **333**, 457–461 (1988).
- Hirsch, V. M., Olmsted, R. A., Murphy-Corb, M., Purcell, R. H. & Johnson, P. R. *Nature* **339**, 389–392 (1989).
- Tsujimoto, H. *et al.* *Nature* **341**, 539–541 (1989).
- Daniel, M. D. *et al.* *Int. J. Cancer* **41**, 601–608 (1988).
- Desrosiers, R. C., Daniel, M. D. & Li, Y. *AIDS Res. Hum. Retroviruses* **5**, 465–473 (1989).
- Peeters, M. *et al.* *AIDS* **3**, 625–630 (1989).
- Roizman, B. *et al.* *Intervirology* **16**, 201–217 (1981).
- Desrosiers, R. C. & Falk, L. A. *J. gen. Virol.* **56**, 119–130 (1981).
- Li, Y., Naidu, Y. M., Daniel, M. D. & Desrosiers, R. C. *J. Virol.* **63**, 1800–1802 (1989).
- Gilden, R. V. *et al.* *Lancet* **i**, 678–679 (1986).

In heterogeneous catalysis, surface-adsorbed molecules are generally the principal reactive species. But the intermediates of such chemical transformations (free radicals, for example) are notoriously difficult to detect because they exist for only a short length of time, and their concentration at any instant therefore tends to be very low. Reid *et al.* now report on page 328¹ a new method for studying possible radical intermediates on solid surfaces. These experiments involve directing a beam of very-high-energy positive muons (μ^+) through the sample — in this case, silica spheres with a monolayer of benzene ($\text{SiO}_2\text{-C}_6\text{H}_6$) — and studying its chemical effects by detecting the positrons that are formed by muon decay about 2×10^{-6} s after the exposure. The radical detected was the muonated cyclohexadienyl radical — as expected, because it is also formed when pure benzene is exposed to muons. Reid *et al.*'s exotic technique allows this radical to be detected despite its short lifetime and its very low concentration. Also, the variations of line-width of the muon spectrum with temperature give information about the motion of these radicals.

Positive muons can be thought of as light protons, the mass being about one-ninth that of the proton. From a chemist's point of view, they are a fortunate product of some of the huge and costly experiments indulged in by physicists in which elementary particles are accelerated and induced to interact with each other. When the muon beam impinges on a collection of molecules, radiation chemistry occurs in the same way as, for example, it does on exposure to high-energy protons, electrons or γ -rays (see

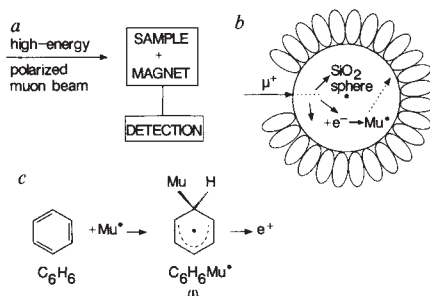
figure). In the $\text{SiO}_2\text{-C}_6\text{H}_6$ experiment, high-energy μ^+ particles will penetrate the system, the most probable reaction resulting in the formation of muonium atoms Mu (this is the equivalent of a hydrogen atom), by electron capture. Of course, other chemical processes take place, but these are not seen by the detection system, which is concerned only with the muons and the associated decay positrons.

In the absence of benzene, normal Mu atoms are detected in the silica grains — their properties correspond exactly to those of hydrogen atoms trapped in silica. For the small size of silica grains used, there is a high probability that a muonium atom generated in the grain can migrate to the surface and react with benzene in the limited time available. Thus the benzene can 'harvest' muonium atoms.

It can be deduced that the $\text{C}_6\text{H}_6\text{Mu}^{\bullet}$ radicals are not free to rotate but that they migrate around the silica spheres in an unrestricted manner. This averages out the anisotropies of the muon–electron–nuclear interactions almost as effectively as free rotation.

How do these results compare with those obtained for benzene molecules using techniques such as infrared or NMR spectroscopy? Extensive studies^{2,3} have shown that these molecules are equally free to migrate but not to rotate freely, the rates deduced in this way being comparable with those for $\text{C}_6\text{H}_6\text{Mu}^{\bullet}$. In hindsight, it is not surprising that this should be the case, because the mass difference is negligible, and there is no specific mechanism for tying the radicals to surface groups on the spheres. Hence they are carried around with the benzene molecules.

Muon spin resonance can be most readily compared with electron spin resonance (ESR) spectroscopy⁴. The main difference between the two techniques is that the first is concerned with coupling to the muon nucleus in radicals whereas the second looks at electrons in radicals. In most cases essentially the same species are detected, and it is the unique isotope effects that are of primary interest. There are, however, some areas in which muon studies are clearly superior to ESR studies. One is in the detection of gas-phase radicals, and another is in the study of various solids, in particular industrially important materials such as crystalline and amorphous silicon, germanium and III–V (such as gallium arsenide) and II–VI (such as cadmium sulphide) compounds. Normal ESR methods are very limited in their ability to detect gas-phase radicals⁵ (mainly because of rota-



a, A muon beam is isolated from the accelerator and impinges on the sample which is housed between the poles of an electromagnet and the positron detection plates. b, Sketch of an individual silica grain which captures a muon to give a muonium atom (Mu). Damage caused by the beam is not detected by this technique. The Mu migrates to the surface and reacts with a benzene molecule to give a cyclohexadienyl radical, as in c, which then emits a positron that is detected by the very sensitive detection system.

tional coupling and consequent line broadening), whereas muonated radicals can be easily detected in the gas phase. Muonated centres have been detected whose hydrogen-atom analogues have never been detected by ESR spectroscopy^{6,7}. Moreover, because of their greater sensitivity, muons can be used to study the kinetics of surface reactions. Perhaps surface studies such as that described by Reid *et al.* will prove to be yet another area in which the use of muons is advantageous, although it will never be a popular method because of the extreme difficulties in obtaining muon beam time. And it must be borne in mind that the muon technique is quite limited in the types of radicals that can be detected, in contrast with ESR spectroscopy, which is also a powerful tool for studying radicals on surfaces. It seems likely, however, that

those who do have access to muons will be inspired by the results of Reid *et al.* to try to probe surface reactions of industrial importance, particularly from a kinetic viewpoint. □

Martyn C. R. Symons is in the Department of Chemistry at the University of Leicester, University Road, Leicester LE1 7RH, UK.

1. Reid, I. D., Azuma, T. & Roduner, E. *Nature* **345**, 328–330 (1990).
2. Kiselev, A. V. & Lygin, V. I. *Infrared Spectra of Surface Compounds* (Wiley, New York, 1975).
3. Boddenberg, B., Haul, R. & Opperman, G. *Naturwiss.* **56**, 636–640 (1969).
4. Cox, S. F. J. & Symons, M. C. R. *Hyperfine Interactions* **32**, 689–706 (1986).
5. Carrington, A. *Microwave Spectroscopy of Free Radicals* (Academic, London, 1974).
6. Patterson, B. D. *Muons and Pions in Materials Research* (eds Chappert, J. & Grynszpan, R. I.) ch. 10 (North Holland, Amsterdam, 1984).
7. Cox, S. F. J. & Symons, M. C. R. *Chem. Phys. Lett.* **126**, 516–525 (1986).

EARTHQUAKES

Seismic cycle not so simple

Max Wyss

UNTIL now it has seemed reasonable to model the seismic cycle of great earthquakes along plate boundaries as follows: at the time of a great earthquake a rupture with approximately uniform slip releases most of the available energy throughout the extent of the rupture outlined by the aftershocks. Then, plate motions accumulate strain energy at an approximately constant rate over the recurrence period (thirty to several hundred years, depending on the location), until the strain level at which rupture occurs is again reached. At that time a great earthquake re-ruptures the same plate boundary segment. This process would result in a seismic cycle of approximately constant recurrence time, because the loading rate and rupture dimensions are constant. The rupture lengths may repeatedly be the same because some geometric tectonic features may define segments of the plate boundary that will rupture in one earthquake. By showing that earthquakes that re-rupture a particular segment of plate boundaries are not similar, Thatcher^{1,2} (first in *Nature* and now in full in the *Journal of Geophysical Research*) has deprived seismologists of part of their favourite model for the generation of great earthquakes.

In cases that conform closely to this model, one speaks of a characteristic earthquake³. There are, however, also sequences of earthquakes in nearly the same location with regular recurrence times but significantly different magnitudes⁴. Thatcher points out that the simple recurrence model is based partly on the incorrect assumption that the slip is evenly distributed over the rupture area, when in reality most of the energy release comes from a fraction of the rupture area, where

the slip is large. He also shows that characteristic earthquakes are the exception rather than the rule along the circum-Pacific plate boundaries. Segments that have historically ruptured in one great earthquake often re-rupture in two or three smaller events, a few years apart. But did these subsequent ruptures really achieve the same strain release as the large one?

As Thatcher observes, the sum of the seismic moment release in several smaller events is much smaller (in the best documented case, 80 per cent less) than that of a single large event. Thus, one cannot think of the two types of rupture as equivalent, even though large events that are multiple ruptures might be interpreted as a sequence of smaller earthquakes occurring within minutes rather than within decades. The slip in the smaller earthquakes, which make up one part of the cycle, may be as little as 20 per cent of the slip in the previous cycle (the large single event), or 40 per cent over half the width of the plate boundary. One might hesitate to call something a 'seismic cycle' which is defined by subsequent episodes of energy release of 100 per cent and 20 per cent. It would be more correct to argue that in the second instance nothing has happened (no energy release) to a first approximation, than to take the view that the same cyclic event (100 per cent release) has happened again.

Thatcher retains the idea of a seismic cycle because the entire plate boundary segment is ruptured in great earthquakes each time, and the intervals remain the same. But do they? Thatcher argues that they do. But this idea may also have to be modified when developing a more compli-

cated model. In the several cases for which Thatcher gives more than one recurrence time (at least three repeated ruptures) the differences between the shortest and longest recurrence interval is on average 67 per cent of the shortest interval (16 per cent and 207 per cent for the extreme cases). Examples of even greater variation of recurrence interval (40 to 300 years) for large ruptures with similar amount of slip exist for segments of the San Andreas fault⁵. Perhaps the whole simple model of recurrence of great earthquakes may have to be substantially revised because the energy release in re-ruptures and the recurrence time vary greatly.

The main value of Thatcher's work is that he has pointed out the inadequacy of a simple model of recurrence that was easily understood and seemed to make sense. The more complicated model he has proposed^{1,2} for explaining the varied nature of re-rupture may not convince all readers. He has not considered the possibility of asperities (strong spots at the fault surface) changing their role from barriers to rupture initiation points as a function of the strain accumulated in them. Also, the amount of slip variation in one earthquake along the fault is not discussed quantitatively, leaving the reader wondering whether a 30 per cent difference in slip is enough to separate asperities from weak fault segments, and how much aseismic fault creep is necessary in Thatcher's model to balance the long-term plate motion along the boundary.

It is not particularly comfortable when 'well established', simple and sensible models are discredited. Thatcher has done this and raised additional questions, which may lead to more changes of the recurrence model of great earthquakes. As Thatcher emphasizes, the role of asperities^{6–10}, their distribution and their activation in subsequent cycles must become clear before we can fully understand the rupture process along great faults and make substantial progress in predicting earthquakes. □

Max Wyss is in the Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, Campus Box 216, Boulder, Colorado 80309-0216, USA.

1. Thatcher, W. *Nature* **341**, 432–434 (1989).
2. Thatcher, W. *J. geophys. Res.* **95**, 2609–2623 (1990).
3. Bakun, W. H. & McEvilly, T. V. *J. geophys. Res.* **89**, 3051–3058 (1984).
4. Wyss, M. *Science* **234**, 726–728 (1986).
5. Sieh, K., Stuiver, M. & Brillinger, D. *J. geophys. Res.* **94**, 603–623 (1989).
6. Wesson, R. L., Burford, R. O. & Ellsworth, W. L. *Proc. Conf. Tectonic Problems of the San Andreas Fault System* (eds Kovach, R. L. & Nur, A.) 303–321 (Stanford University, Stanford, 1973).
7. Aki, K. *J. geophys. Res.* **84**, 6140–6148 (1979).
8. Lay, T. & Kanamori, H. in *Earthquake Prediction* (eds Simpson, D. W. & Richards, P. G.) 579–592 (Am. Geophys. Union, Washington, D.C., 1981).
9. Wyss, M., Johnston, A. C. & Klein, F. W. *Nature* **289**, 231–234 (1981).
10. Aki, K. *J. geophys. Res.* **89**, 5867–5872 (1984).

Some relatives take a dive

R. D. Martin

BECAUSE of their very obscurity, the earliest origins of modern primates remain a topic rooted more in conjecture than in fact. The reports by Beard¹ and by Kay *et al.*² on pages 340 and 342 of this issue, concluding that one family of so-called 'archaic primates' can be ruled out of line, provide much-needed evidence and at the same time give cause for a fresh round of speculation.

By mid-Eocene times, 45 million years ago, prosimian primates of modern aspect were well established, being documented by associated skulls and skeletons from both North America (for example *Notharctus*³) and Europe (*Euro-polemur*⁴ and *Pronycticebus*⁵). By this stage, many characters typical of modern primates — large, forward-facing eyes surrounded by a complete bony rim, hindlimb domination and grasping feet — had developed and a great diversity of species was already present. Indeed, evidence from the Fayum site (Egypt) indicates that even advanced simian primates had evolved by the end of the Eocene⁶. Hence, the common ancestor of modern primates undoubtedly dates back well into the Palaeocene, some 60 million years ago, if not earlier; but Palaeocene primates of modern aspect have yet to be documented. Instead, Palaeocene deposits of North America and Europe have yielded four families of so-called 'archaic primates' (Plesiadapidae, Paromomyidae, Picrodontidae, Saxonellidae), which are usually combined in the infraorder Plesiadapiformes of the order Primates⁷⁻⁹. These fossils resemble later primates in certain features of the molar teeth and the ear region, but lack many other typical primate characters. Some authors have therefore questioned the inferred phylogenetic link between plesiadapiforms and primates¹⁰⁻¹².

So far, this issue has remained contentious because only the single genus *Plesiadapis* was known from reasonably well-preserved skulls and skeletons^{8,9,13}. The picture has now radically changed following the discovery of new material from two late-surviving genera of the family Paromomyidae in earliest Eocene deposits of Wyoming. Here Beard¹ reports on two partial skeletons of *Phenacolemur*, and Kay and colleagues² describe a virtually complete skull of *Ignacius*. Both Beard and Kay *et al.* reach

the same revolutionary conclusion — that paromomyids are related not to primates but to the Dermoptera, a peculiar order of Southeast Asian gliding mammals containing only two species (genus *Cynocephalus*).

The partial skeletons discussed by Beard were associated with dental remains, permitting them to be identified



Lithograph of a living 'flying lemur' (Dermoptera) from *Histoire Naturelle des Singes et des Makis*, 1800. The plate shows the animal's elongated middle digits, also found by Beard in his fossil specimens, which are necessitated by the extension of the gliding membrane on to the hands and feet. (Although dermopterans are commonly referred to as flying lemurs, as George Gaylord Simpson pointed out they are not lemurs and they do not fly!)

with confidence as *Phenacolemur*. As a spin-off, isolated postcranial elements could be allocated to *Ignacius* and *Phenacolemur*. In both, the middle phalanges (digital bones) of the fingers and toes are markedly longer than the basal phalanges. This unusual condition is lacking among primates, living or fossil. Among modern mammals, dermopterans provide the only counterpart: the gliding membrane (patagium) extends onto the digits of the hands and feet, which explains the elongated middle phalanges. Beard also identified derived features of the wrist and ankle, which suggest a link between *Phenacolemur*, *Ignacius* and *Cynocephalus*, and he concludes that Paromomyidae and

Dermoptera are sister groups.

Kay *et al.* show that the skull of *Ignacius* shares certain derived features with dermopterans and differs distinctly from primates. One diagnostic feature of primates of modern aspect is formation of the auditory bulla primarily as an outgrowth of the petrosal bone^{8,9,12}. In *Ignacius*, by contrast, the bulla is apparently formed by a separate element, an entotympanic. Re-examination of the evidence for *Plesiadapis* indicates that in this genus, too, the bulla is formed by an entotympanic. The inferred presence of a petrosal bulla in *Plesiadapis* has provided one of the main arguments for a phylogenetic link between plesiadapiforms and primates. But it is notoriously difficult to determine the composition of the bulla in adult mammals because of fusion of sutures¹⁴. (Tree-shrews, once thought to have a petrosal bulla, provide a classic example — from developmental evidence, it eventually became clear that the bulla is formed from an entotympanic that later fuses with the petrosal.)

Kay *et al.* also infer that *Plesiadapis*, *Ignacius* and *Cynocephalus* share an unusual derived feature in that the internal carotid system is only poorly developed, the principal intracranial blood supply being provided through the vertebral arteries. They therefore go further than Beard and propose that the entire infraorder Plesiadapiformes represents the sister-group of the Dermoptera.

The new evidence opens up several possibilities. The first is that Dermoptera and Primates are related and that the Plesiadapiformes represent an important connecting link between them. This would fit the idea, now experiencing a renaissance, that tree-shrews, bats, dermopterans and primates constitute a phylogenetic unit (the 'Archonta'). On the other hand, resemblances between these

four groups may simply reflect retention of primitive features from ancestral placental mammals¹⁴. A second possibility (left open by Beard, but ruled out by Kay *et al.*) is that plesiadapids are related to primates of modern aspect, while paromomyids are related to dermopterans. But the evidence for a link between plesiadapids and primates is considerably weakened if the former have an entotympanic bulla.

Third, as Kay *et al.* suggest, the entire infraorder Plesiadapiformes may be specifically related to dermopterans. Evidence from the intracranial blood supply supports this interpretation, but there is a problem with the bulla — Kay

et al. cite fusion of a collar-shaped ectotympanic to the outside of the bulla as a shared derived feature of plesiadapiforms and dermopterans, yet Gingerich⁹ has described a well-preserved skull of *Plesiadapis* that clearly shows an ectotympanic ring enclosed within the bulla. Finally, it could be that plesiadapiforms are connected neither with primates nor with dermopterans, and that the paromomyids convergently developed gliding adaptations paralleling those of modern dermopterans.

As dental evidence has figured so

prominently in discussions of the purported link between plesiadapiforms and primates, it is curious that no specific dental resemblances between *Cynocephalus* and plesiadapiforms have been proposed. Whatever the outcome, the new evidence challenges the long-accepted relationship between plesiadapiforms and primates, and it shows that paromomyids, at least, were sky-divers. □

R. D. Martin is in the Anthropologisches Institut und Museum, Universität Zürich-Irchel, 8057 Zürich, Switzerland.

1. Beard, C. *Nature* **345**, 340–341 (1990).
2. Kay, R., Thorington, R. & Houde, P. *Nature* **345**, 342–344 (1990).
3. Gregory, W. K. *Mem. Am. Mus. nat. Hist.* **35**, 239–257 (1920).
4. Franzen, J. L. *Cour. Forsch.-Inst. Senckenberg* **91**, 151–187 (1987).
5. Thalmann, U., Haubold, H. & Martin, R. D. *Palaeoverbrata* **19**, 115–130 (1989).
6. Simons, E. L. *Science* **247**, 1567–1569 (1990).
7. Simons, E. L. *Primate Evolution: An Introduction to Man's Place in Nature* (Macmillan, New York, 1972).
8. Szalay, F. S. & Delson, E. *Evolutionary History of the Primates* (Academic, New York, 1979).
9. Gingerich, P. D. *Mus. Pal. Univ. Mich., Pap. Paleont.* **15**, 1–140 (1976).
10. Cartmill, M. in *The Functional and Evolutionary Biology of Primates* (ed. Tuttle, R.) 45–77 (Aldine-Atherton, Chicago, 1972).
11. Martin, R. D. in *Major Topics in Primate and Human Evolution* (eds Wood, B. A., Martin, L. B. & Andrews, P. J.) 1–31 (Cambridge University Press, 1986).
12. Martin, R. D. in *The Study of Prosimian Behaviour* (eds Doyle, G. A. & Martin, R. D.) 45–77 (Academic, New York, 1979).
13. Szalay, F. S., Tattersall, I. & Decker, R. L. *Contrib. Primatol.* **5**, 136–166 (1975).
14. Martin, R. D. *Primate Origins and Evolution: A Phylogenetic Reconstruction* (Chapman & Hall, London/Princeton University Press, 1990).

OXIDE SUPERCONDUCTORS

Contrasting critical currents

Neil Alford

APPLICATIONS of high-temperature superconductors have so far been severely limited because their current densities are disappointingly low. To be useful, these materials need to have zero resistance at liquid-nitrogen temperatures or higher, with large current densities in high magnetic fields, and the samples need to be large (of the order of metres) for use as wires or cables, for example. On page 326, Ru Ling Meng *et al.*¹ describe a 'continuous' process for making $\text{YBa}_2\text{Cu}_3\text{O}_x$ with high current densities. The process is continuous in the sense that the sample is moved through a hot zone in a furnace and directional solidification is encouraged. In principle, such a process may enable samples of large dimensions with controlled grain orientation to be made, although, as yet, the samples are only centimetres in length.

There are three important factors that limit the current-carrying capacity of the bulk high-temperature superconductors: the presence of weak links at grain boundaries; the anisotropy in the current-carrying ability in crystals of the material — best in the a - b plane and poor in the c -axis direction; and the weakness of flux pinning. In thin films, experiments at IBM² have shown that although the critical-current density, J_c , within a single grain can be of the order of millions of amps per cm^2 , the inclusion of a grain boundary and misorientation between crystals can cause a drastic reduction in J_c of about two orders of magnitude.

The first to achieve high critical-current

density, J_c , in bulk $\text{YBa}_2\text{Cu}_3\text{O}_x$ (123) was Sungho Jin at Bell Labs³, who showed that, by growing large single crystals, J_c could be increased from the usual $1,000 \text{ A cm}^{-2}$ to over $17,000 \text{ A cm}^{-2}$ (at temperatures of 77 K in zero magnetic field). More importantly, the J_c of such materials was less affected by magnetic fields than conventionally sintered, unoriented material. Murakami *et al.*⁴ then showed that by melting, quenching and then annealing the $\text{YBa}_2\text{Cu}_3\text{O}_x$, exceptionally high J_c could be achieved even in very high magnetic fields of up to several tesla. Salama *et al.*⁵ at Houston have confirmed these results and obtained a current density of $37,000 \text{ A cm}^{-2}$ in a magnetic field of 0.6 T by direct (pulsed) measurement. A number of other groups (see, for example, refs 6, 7 and 8) have now obtained high values of J_c in small samples. All these methods make use of the liquid phase in the peritectic regime. Indeed, Murakami *et al.* believe that the problem of weak flux pinning may be overcome by the presence of the second-phase material which forms as a consequence of these peritectic growth methods. The peritectic region is one where a solid of composition A and a liquid of composition B combine to form a solid on cooling of composition C. The 123 compound will grow from a liquid in this region and may form as a single-phase material, but from experience second-phase particles such as Y_2BaCuO_x (211) form. If the melt is quenched, very fine particles of Y_2O_3 , for example, may also remain. So why is the second-phase

impurity useful?

A perfect crystal of high- T_c superconducting material has zero J_c , the exact opposite of what is required. Only when defects, which will pin the magnetic flux, are introduced will current be carried without the flux lattice disappearing. This is because the Lorentz force between the transport current and the flux lines causes the latter to move, thus generating a voltage, unless the flux lines are pinned. The defects are really the glue which holds the lattice together and without them there can be no transport of current. These defects must be of the order of the coherence length, which in the new oxide superconductors is extremely small — about 3 nm. The process of using a liquid phase in the peritectic region naturally produces second-phase material, and it is possible that by quenching, the dimensions of these second-phase particles can be constrained to the dimensions of the coherence length. Defects can also be introduced by irradiation and groups at Argonne, AT & T Bell Laboratories and Houston have obtained large increases in J_c in this way.

In the 'powder in a single tube' approach, the superconductor of choice is Pb-doped Bi-Sr-Ca-Cu-O (usually the three-layer compound 2223) because of the 'platey', mica-like particle morphology⁹. The swaging (reducing the cross-section by force) and rolling process used is standard metallurgy and the results, $J_c = 32,000 \text{ A cm}^{-2}$ at 77 K in zero field, $23,000 \text{ A cm}^{-2}$ in 0.1 T and $6,000 \text{ A cm}^{-2}$ in 1 T at 77 K are impressive. A uniform response can be maintained over several centimetres, and even at lengths of 2–3 m a J_c of $10,000 \text{ A cm}^{-2}$ at 77 K in zero field can be attained (Ken Sato of Sumitomo Electric, personal communication). The microstructure of these materials shows excellent alignment of the a - b plane and one can imagine that the repeated sintering and rolling steps gives good intergranular contact thereby reducing the weak-link problem. Interesting work by Vacuumschmelze in Germany shows that silver-clad $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_x$ has high J_c which is not affected by the presence of magnetic fields, with properties at 4 K rivaling those of conventional superconductors. Unfortunately, however, at 77 K the properties are poor.

New work at Southampton University is also throwing up some exciting indications of what may be possible¹⁰. This group finds that the transport J_c of melt-processed $\text{YBa}_2\text{Cu}_3\text{O}_x$ is $16,000 \text{ A cm}^{-2}$ at 0 T and 77 K and $2,000 \text{ A cm}^{-2}$ at 1 T and 77 K. What is interesting is that transport a.c. measurements are much higher, $J_c = 12,000 \text{ A cm}^{-2}$ at 5 T and 77 K measured at 50 Hz.

The possibility of continuous processing raised by Meng *et al.*¹ brings practical applications for high- T_c superconductors closer to realization, because of the potential to maintain high J_c over useful lengths.

It also seems likely that such material will have low surface resistance and so it could be useful in the area of radio-frequency and microwave applications.

Meng *et al.* measured a transport J_c at 77 K in magnetic fields and obtained 7,500 A cm⁻² at 0.82 T parallel to the direction of transport and 14,000 A cm⁻² perpendicular to it. The material shows, therefore, far better field dependence than conventionally sintered materials where, at 0.82 T, the J_c is less than 10 A cm⁻². The materials are comparable, or possibly slightly inferior, to the materials made by Murakami or Salama using a melt and quench process — this, however, is a batch, not a continuous process.

What are the prospects? All the work on melt-processed samples mentioned here has been undertaken on what can only be described as tiny samples, and cables kilometres in length with high J_c are still a long way off. Moreover, there are other problems yet to solve. For example, the kinetics of the crystallization process dictate that it is bound to take a long time and then because the material contains no

interpenetrating porosity the oxygenation required could be prohibitively long¹¹. The fracture toughness will still be that of glass so handling may be a problem. But science is generally hard graft interspersed with the occasional breakthrough. We have seen that in the high-temperature superconductors the properties are there for the taking. The paper from Professor Chu's group in this issue is an indication of what may be possible. □

Neil Alford is a Division Research Associate in the ICI Advanced Materials, Solid State Science Group, The Heath, Runcorn, Cheshire WA7 4QE, UK.

1. Meng, R.L., Kinaldis, C., Sun, Y.Y., Gao, L., Tao, Y.K., Hor, P.H. & Chu, C.W. *Nature* **345**, 326–328 (1990).
2. Chaudhari, P. *et al.* *Phys. Rev. Lett.* **60**, 1653 (1988).
3. Jin, S. *et al.* *Phys. Rev.* **B37**, 7850 (1988).
4. Murakami, M. *et al.* *Jap. J. appl. Phys.* **28**, 1189–1194 (1989).
5. Salama, K. *et al.* *Appl. Phys. Lett.* **54**, 2352–2354 (1989).
6. McGinn, P.J. *et al.* *Physica* **C156**, 57–61 (1988).
7. Alford, N. *et al.* *J. appl. Phys.* **65**, 2856–2858 (1989).
8. Pollard, R.J. *et al.* *Supercond. Sci. Technol.* **2**, 169–172 (1989).
9. Hikata, T. *et al.* *Jap. J. appl. Phys.* **28**, L82–L84 (1989).
10. Yang, Y. *et al.* *Supercond. Sci. Technol.* (in the press).
11. Alford, N. *et al.* *Nature* **332**, 58–59 (1988).

ELECTROCHEMISTRY

More efficient solar cells

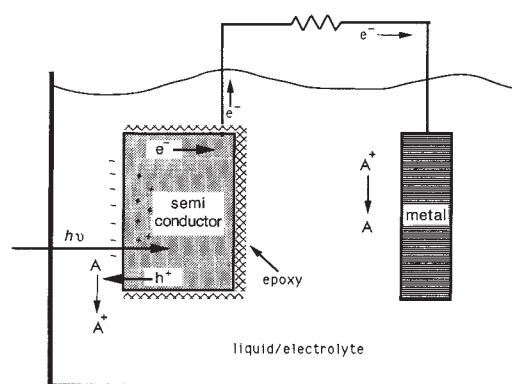
Nathan S. Lewis

AS THE search continues for an inexpensive, efficient solar-energy-conversion device, conventional photovoltaic devices are under increasing competition from 'wet' photoelectrochemical cells. Photoelectrochemical cells are relatively simple (and potentially inexpensive) to construct, because they require merely the immersion of the desired photoactive solid into an ionically conducting liquid. On page 330, a team led by S. Licht¹ describe how in certain cases it is possible to achieve the desirable combination of high efficiency and low corrosion rates by manipulating the composition of the conducting liquid. With their approach, Licht *et al.* demonstrate that cadmium selenide electrodes can be coaxed into stable operation in an aqueous solution with an overall energy-conversion efficiency of 17 per cent.

In the general scheme of energy-conversion technologies, semiconductor/liquid cells fall in between solid-state photovoltaics and artificial photosynthetic systems. The solid-state systems are the most advanced, efficient and reliable devices, with some specialized laboratory devices yielding terrestrial efficiencies in the 25–26 per cent range. But their efficiency comes at a price, both in terms of the expense in making the material and the complex processing steps used to coax out the last percentage points in efficiency. At present, semiconductor/liquid systems offer efficiencies in the 10–15 per cent range, with demonstrated lifetimes in

several systems now ranging from months to years. The molecular-based artificial photosynthetic systems are the least efficient and stable, with efficiencies on the order of one per cent and stability limited by molecular photodegradation processes. But the molecular systems do at least offer an opportunity for chemical design of the entire system, which might eventually enable construction of a photoresponsive assembly that is analogous to the tried-and-true method of natural photosynthesis.

The principles leading to photocurrent flow at a semiconductor/liquid junction, like that used by Licht *et al.*, are very similar to those of conventional photovoltaic cells (see figure). The junction,



Schematic of a photoelectrochemical cell. Light absorbed by the semiconductor electrode is converted into a flow of electrical charge in the external circuit.

between the solid and liquid produces an electric field, which separates photo-generated charge carriers, electrons and holes. Carrier recombination can be minimized, and efficient chemical energy conversion is possible, if these charge carriers can be transferred across the solid/liquid interface and onto the desired ions in the liquid. Manipulation of the reactions that occur at the two electrodes in the photoelectrochemical cell can result in the production of electricity, fuels or both².

One of the key problems plaguing the development of these systems is the trade-off between the stability of the solid and the efficiency of the resulting device: semiconductors that are stable in the presence of light often make inefficient photovoltaic devices, whereas the alternative solid/liquid combinations generally corrode when illuminated. Several strategies are being used to tackle this problem — non-aqueous solvents can be used to suppress corrosion; novel semiconductors are being developed that are more photochemically inert in aqueous solutions; the surface of traditional semiconductors is being modified chemically; and manipulation of the composition of the ions in the aqueous electrolyte is being investigated. All these methods have particular advantages, but it is not yet clear which strategy is most effective.

In Licht's work, the authors have used cadmium selenide as the photoanode in an aqueous cell containing hexacyanide complexes of Fe(II) and Fe(III) as the charge carriers. They find that at specific values, of the Fe(II) and the Fe(III) concentrations, the photoanode is stable under illumination for greater than 5,000 Coulombs cm⁻² of total charge flow. Specific concentrations are required to achieve a fine balance between the diffusive flux of ions to the surface, the need rapidly to capture the photogenerated charges before they have a chance to participate in corrosion processes, and the tendency to deposit insulating layers of chemical films if the solubility of the reaction products is exceeded. Achieving this delicate balance in the aqueous cadmium selenide system has allowed Licht *et al.* to improve the cell stability from seconds to days, while simultaneously achieving excellent output properties.

The magnitude of the voltage developed in these cells is possibly even more impressive than their improved stability. Many workers associate the performance of semiconductor/liquid junctions with that of semiconductor/metal (Schottky) contacts. Unfortunately, Schottky contacts are notoriously inefficient photovoltaic devices, because the metal overlayer efficiently collects both electrons and holes. This leads

to excessive recombination as compared to p-n junction-type photovoltaic devices. Early work indicated that excessive recombination was also prevalent at solid/liquid interfaces, but more recent studies have shown that junction recombination can be greatly suppressed with the proper choice of semiconductors and redox ions. In fact studies of silicon/liquid interfaces have shown that extremely high open-circuit voltages can be obtained, even by comparison to well-developed silicon solid-state photovoltaic cells¹. A similar situation appears to be occurring in the cells of Licht *et al.*. Their cells develop photopotentials of 1.2 V, and the band-gap energy is 1.7 electron volts (eV). This ratio is very close to the upper limit that could be achieved from a material with a 1.7-eV band gap, again indicating that near-optimal performance can be achieved in certain semiconductor/liquid combinations. The ability to reach this voltage is even more impressive given the constraints on corrosion suppression, film formation and interfacial kinetics that must be addressed simultaneously in the cadmium selenide system.

One of the unknown features of this system is the rate at which charge carriers actually cross the interface. An understanding of these rates is important, because rapid capture of carriers will allow effective competition with both recombination and corrosion. The few kinetic measurements that have been performed to date on semiconductor/liquid junctions have yielded a wide disparity in the timescales for the charge-transfer process, with values ranging from picoseconds to milliseconds. The factors that control these rates are not well understood at present, and extensive research is still required in order to obtain interfaces with the desired kinetic properties.

In contrast to extraction of energy from fossil fuels or nuclear power, active uses of solar energy are unique in that they do not directly contribute to global warming. As global warming issues gain the attention of the public and policymakers, and more attention is paid to evaluating the true cost of fossil fuels, solar energy processes should regain lost attention when considering alternative energy options. Whether the specific solar technology involves semiconductor/liquid junctions, solid-state systems, artificial photosynthetic devices or some hybrid thereof remains to be seen. □

Nathan S. Lewis is in the Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA.

Following proteins in time

Louise N. Johnson

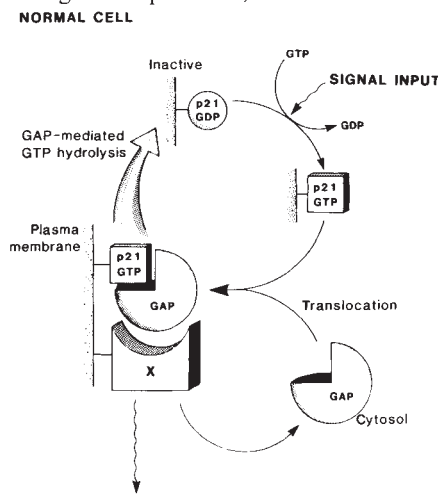
UNLIKE conventional X-ray crystallography, Laue protein crystallography, which uses the bright, broad spectrum of X-radiation emitted by synchrotron radiation sources, has great potential for time-resolved studies on the dynamic properties of proteins in the crystalline state. In 1984 it was demonstrated that the method (white X-ray beam: stationary crystal) was feasible for proteins^{1,2}, and in 1987 that it could be applied to elucidate ligand binding³. Now, on page 309 of this issue, Schlichting *et al.*⁴ show that the method can provide time-resolved information about a biologically important molecule. The authors report a detailed structural study using the DESY synchrotron, Hamburg, on the active Ras p21-GTP complex and a complex on the pathway to the product p21-GDP complex. The hydrolysis of GTP proceeds with a slow half-life (about 40 min in the crystal, similar to that in solution) and the first image of the enzyme-GTP complex was obtained 4 min after starting the reaction with a total exposure of 45 s. The p21 crystals are sensitive to radiation damage; Schlichting *et al.* obtained only two data sets for each crystal, and only one in ten crystals was sufficiently well ordered. The method is not without its technical difficulties.

The advances in Laue crystallography⁵ have depended on advances in physics and in chemistry. The physics has been necessary not only in the design and operation of synchrotron radiation sources and protein crystallography stations, but also for the development of optimum protocols and software to obtain reliable intensity estimates from the diffraction photographs. Advances in chemistry, in particular the development of caged compounds^{6,7}, have allowed the

start of the reaction to be synchronized with the start of the collection of X-ray data. Here, the substrate is made biologically inert through covalent attachment of a photolabile protecting group (Schlichting *et al.* used a 2-nitrophenyl ethyl group). Illumination by a xenon flash lamp results in dissociation of the protecting group and liberation of the substrate. Such strategies have been employed previously in the study of muscle ATP hydrolysis and neurotransmitter mechanisms.

Schlichting *et al.* co-crystallized the p21 protein in the presence of caged GTP and elucidated the structure by conventional data collection methods. After illumination (ten flashes from a xenon flash lamp over about 1 min), the authors studied the resulting active GTP complex some 4 min and 14 min after reaction by Laue crystallography and after 19 h by conventional methods. Even though only 50 per cent of the data to 2.8-Å resolution were obtained with the Laue method, the electron-density maps and the crystallographic refinement of the structures were sufficiently well-defined to give precise information about the complexes. In agreement with analysis of the complexes using high-performance liquid chromatography⁸, the image after 4 min was compatible with a GTP complex, that after 14 min with a mixed state in which there had been hydrolysis, and that after 19 h was consistent with the GDP complex. Interestingly, the caged GTP molecule, bound in a nonproductive mode, was able to rearrange in the crystal after liberation of the cage to yield the active GTP complex.

What does this teach us about the biological function of the Ras p21 protein? The family of *ras* genes in the mammalian genome has been intensely studied following the identification of their transformed alleles in about 10 per cent of human cancer tumours⁹. The *ras* genes acquire their transforming properties by single point mutations in their coding sequences corresponding to positions 12, 13, 59 and 61 of the expressed protein p21. The cellular role of the protein is not known, but by analogy to the G proteins in signal transduction pathways and the bacterial elongation factor Tu, p21 is thought to function by exchanging GTP for GDP in response to an extracellular signal. The p21-GTP complex is then recognized by the GTPase-activating protein (GAP) and the resulting complex acts as a signal for growth¹⁰ (see figure). In the normal cell, the signal is switched off by p21-catalysed hydrolysis of GTP, the catalytic activity being greatly



A scheme for p21 activity and function (from ref. 10, with permission).

1. Licht, S. *et al.* *Nature* **345**, 330–333 (1990).
2. Photocatalysis: Fundamentals and Applications, (eds Serpone, N. & Pelizzetti, E.) (Wiley, New York, 1989).
3. Semiconductor Electrodes (ed. Finkles, H.O.) (Elsevier, Amsterdam, 1988).

increased by interaction with GAP. In the mutant proteins, GTPase activity is low and is not increased by GAP. So the protein remains in the 'on' state, prolonging a signal for unregulated cell growth.

Precise information about the structural changes has been provided by the Berkeley group for GDP and inactive GTP analogue complexes¹¹, by the Heidelberg group¹² for the refined structure of an inactive GTP analogue complex (now at 1.35-Å resolution¹³), and by the new work⁴ for active GTP and GDP complexes. The recombinant p21 proteins differ slightly — residues 1–166 (Heidelberg); residues 1–177 (Berkeley) — and both lack the C-terminal residues involved in attachment of the protein to the cell membrane. The proteins crystallize in different space groups and the crystallographic analysis therefore provides several independent views of the molecule.

Although a detailed comparison of the proteins has still to be made, some common features are apparent. Both groups describe a similar constellation of the amino acids that contribute to the binding of the base, ribose and α -phosphate components of the guanine nucleotides and there is negligible change in these between the GTP and GDP complexes. In the GTP complexes, the β - and γ -phosphates are doubly coordinated by the magnesium ion and by the NH_3^+ group of Lys 16. In addition, the main-chain NH of Thr 35 and Gly 60 hydrogen-bond to the γ -phosphate.

There are differences in detail that may be important. Schlichting *et al.*⁴ report an interaction between Gln 61 and the γ -phosphate through a water molecule, and in the active GTP complex there is an additional hydrogen bond between the β - γ bridging oxygen and the main-chain NH of Gly 13. (GTP has 10–100 times the affinity of the inactive analogue for p21). The main changes in structure between the p21·GTP complexes and the p21·GDP complex occur in two loop regions, L2 (movements of residues 32–

38) and L4 (movements of residues 60–65). Both regions are involved in the recognition of the γ -phosphate. Residues 32–36 have been implicated in GAP recognition by mutational experiments¹⁴. The catalytic mechanism, which involves attack by a water molecule¹⁵, is not yet known, but is likely to involve Lys 16 in transition-state stabilization, Gln 61 in the attack by water, and Thr 35 and Asp 57 in the magnesium and water coordination¹³. At present there is no outstanding candidate for promoting activation of the water, unlike (for example) phosphofructokinase, where the role of Asp 127 is clear¹⁶.

Why are the mutant proteins not activ-

ated by GAP and why do they have negligible GTPase activity? Surprisingly, comparison of the p21 mutant (Gly 12 replaced with Val) and native GTP complexes shows similar binding modes of GTP and only small differences near the loop 10 to 18 that wraps around the phosphate-recognition site. There are differences, however, in conformation of residues Gly 60, Thr 38 and Gln 61 that are involved in the γ -phosphate recognition site, and these changes are likely to affect catalysis. □

Louise N. Johnson is in the Laboratory of Molecular Biophysics, University of Oxford, Oxford OX1 3QU, UK.

BIOGEOCHEMICAL CYCLES

El Niño and atmospheric CO₂

Ulrich Siegenthaler

ON average, the amount of carbon dioxide in the atmosphere is steadily increasing by about 1.5 parts per million (p.p.m.) per year. The increments from one year to the next are, however, not constant but fluctuate around a mean trend. These fluctuations are associated with El Niño/Southern Oscillation (ENSO) events. During an ENSO event, the sea surface temperatures in the tropical Pacific Ocean increase and the upwelling of nutrient-rich water ceases, thereby increasing the uptake of carbon dioxide from the atmosphere by the ocean. The apparent link between ENSO events and fluctuations in atmospheric carbon dioxide concentrations seems plausible. The simplicity of this argument is, however, misleading. C. D. Keeling and others¹ have carefully analysed measurements of the concentration and ¹³C/¹²C ratio ($\delta^{13}\text{C}$ value) of atmospheric carbon dioxide and find that biospheric, as opposed to oceanic, processes are the primary cause of the interannual fluctuations in atmospheric carbon dioxide.

To study the interannual fluctuations in CO₂, Keeling *et al.* first filtered out the seasonal cycle of CO₂ from the records, then they removed the long-term increas-

ing trend by subtracting a constant fraction of the cumulative fossil CO₂ emissions. Figure 1 shows the average of the resulting anomalies in CO₂ concentrations over Mauna Loa and the South Pole; this curve shows, to a first approximation, fluctuations of the global mean atmospheric CO₂ content. The correlation of the CO₂ anomalies with ENSO events is obvious: during such an event, a rise in CO₂, exceeding the normal anthropogenic growth rate, starts and lasts for about a year. Before such an increase, a dip is sometimes observed. The interannual CO₂ variations must reflect imbalances in the exchange fluxes between the atmosphere and the ocean and/or the terrestrial biosphere. Although the oceanic and biospheric sources of CO₂ cannot of course be distinguished by examining the concentration data alone, they can be distinguished by looking at the $\delta^{13}\text{C}$ value of the atmospheric CO₂. Carbon dioxide released from the ocean has nearly the same $\delta^{13}\text{C}$ as atmospheric CO₂ (about -8‰), because atmospheric CO₂ is in approximate isotopic equilibrium with dissolved inorganic carbon in sea water. In contrast, CO₂ of biospheric origin typically has a $\delta^{13}\text{C}$ of

1. Moffat, K., Szebenyi, M.E. & Bildereck, D.H. *Science* **223**, 1423–1425 (1984).
2. Helliwell, J.R. *Rep. Prog. Phys.* **47**, 1403–1497 (1984).
3. Hadju, J. *et al. Nature* **329**, 115–116 (1987).
4. Schlichting, I. *et al. Nature* **345**, 309–315 (1990).
5. Hajdu, J. & Johnson, L.N. *Biochemistry* **29**, 1669–1678 (1990).
6. Kaplan, J.H., Forbush, B. & Hoffman, J.F. *Biochemistry* **17**, 1929–1935 (1978).
7. McCray, J.A. & Trentham, D.R. *A. Rev. Biophys. Chem.* **18**, 239–270 (1989).
8. Schlichting, I. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 7687–7690 (1989).
9. Barbaud, M. *A. Rev. Biochem.* **56**, 779–827 (1987).
10. McCormick, F. *Cell* **56**, 5–8 (1989).
11. Milburn, M.V. *et al. Science* **247**, 939–945 (1990).
12. Pai, E.M. *et al. Nature* **341**, 209–214 (1989).
13. Pai, E.M. *et al. EMBO J.* (in the press).
14. Gibbs, J. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 5026–5030 (1988).
15. Feuerstein, J., Goody, R.S. & Webb, M.R. *J. biol. Chem.* **264**, 6188–6190 (1989).
16. Shirakihara, Y. & Evans, P.R. *J. molec. Biol.* **204**, 973–994 (1988).

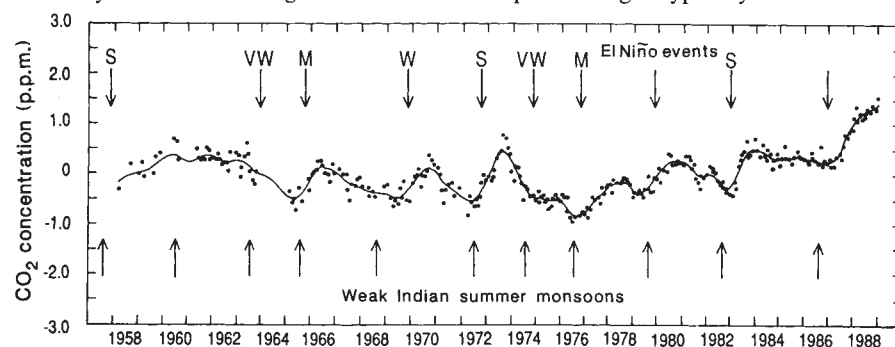


Fig. 1 Atmospheric CO₂ concentration anomaly obtained by Keeling *et al.*¹ by removing the seasonal variations as well as the secular trend from their observational records. The data points represent averages of the anomalies of Mauna Loa and the South Pole. The strength of the El Niño events is indicated, ranging from very weak (VW) to strong (S).

–25‰ and is thus clearly distinguishable from that originating from the ocean.

Figure 2 shows a seasonally adjusted record of $\delta^{13}\text{C}$. Excursions towards lower values of $\delta^{13}\text{C}$ (notice inverted scale) are observed in 1983–84 and 1987–88 and also 1979–80, occurring at the same time as the anomalous increases in CO_2 (Fig. 1) after an ENSO event. The anomalies are small, comparable to the analytical precision, but it seems unlikely that they are an experimental artefact, as they are found in the individual isotopic records for both Mauna Loa and the South Pole. Thus, the isotope data suggest that the positive CO_2 anomalies are of biospheric origin.

Examining both the isotopic records and the records of CO_2 concentration allows both the biospheric and oceanic sources to be identified: from the anomalies in CO_2 concentration, the sum of biospheric and oceanic sources (or sinks) can be estimated, and anomalies in $\delta^{13}\text{C}$ allow the size of the biospheric contribution alone to be established. Keeling *et al.* have carried out a double deconvolution of the two records and obtained estimates for the 'anomalous' oceanic and biospheric CO_2 fluxes. Both fluxes are found to be large, up to 4 Gt C per year, and are out of phase: namely, biospheric release and oceanic uptake of carbon occur during and after an ENSO event, with the biospheric flux dominating, thus causing a net release of CO_2 to the atmosphere.

Is this result realistic? Qualitatively, it is certainly plausible. Ordinarily, surface waters in the equatorial Pacific Ocean have a high partial pressure of CO_2 (P_{CO_2}), with values 50–100 p.p.m. higher than in the atmosphere. This peak in P_{CO_2} is due to the upwelling of CO_2 -rich water from depths of 100 m to a few 100 m and the resulting degassing from the equatorial Pacific corresponds to a source of about 1 Gt C per year. During an El Niño event, upwelling stops and therefore the CO_2 outgassing ceases. Although this is directly supported only by a few P_{CO_2} measurements during the 1982 event, there is clear evidence from data on total carbon and on nutrients that P_{CO_2} must have decreased. This is confirmed by box model studies^{2,3} in which the physical, chemical and biological changes in the Pacific Ocean during an ENSO event are simulated. During such an event, the upwelling of water rich in CO_2 and nutrients into the equatorial surface mixed layer ceases; the temperature of the surface water increases and biological activity decreases (because of reduced availability of nutrients), as does the export of carbon by sinking organic particles. These last two effects would tend to increase P_{CO_2} in the surface water,

but the first — which dominates — leads to a decrease of P_{CO_2} . Although the two model studies^{2,3} differ in the details, both obtain the same decrease, about 0.4 p.p.m. (or 0.8 Gt C), in atmospheric CO_2 during an El Niño. The pattern of oceanic fluxes as predicted by the box models is in qualitative agreement with the pattern deduced by Keeling *et al.*, but their estimate of the oceanic CO_2 flux, up to 2 p.p.m. change within a year, is larger than can be explained by the models.

The biospheric contribution also seems plausible. During an El Niño, rainfall is

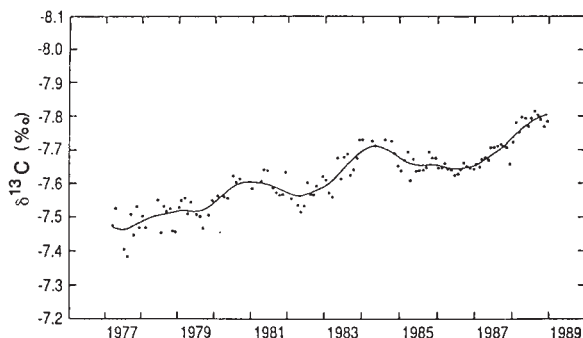


Fig. 2 Seasonally adjusted $\delta^{13}\text{C}$ values in atmospheric CO_2 , average of results from Mauna Loa and the South Pole¹. The vertical scale is inverted to facilitate comparison with Fig. 1.

reduced in many tropical regions; and the correlation between El Niño and the weak southeast Asian summer monsoon is striking (Fig. 1). Consequently, there is apparently reduced uptake of CO_2 by vegetation over large areas, so that the anomalous atmospheric CO_2 increase during an El Niño seems essentially to be due to a reduced biospheric CO_2 sink.

In view of the relatively small CO_2 signals involved, one may wonder whether the story is relevant to discussions on the carbon cycle — the ENSO-related anomalies can hardly lead to changes in atmospheric CO_2 that would influence the greenhouse effect. The observed anomalies of 1–2 Gt C per year do, however, correspond to several per cent of the total terrestrial net primary productivity. This is a significant perturbation, particularly because it has its origin in only a restricted part of the land biota, and is important enough to be studied in more detail. As usual, what are needed are further long-term high-precision records of $\delta^{13}\text{C}$ and CO_2 concentration from stations near tropical ecosystems, and direct P_{CO_2} observations in the equatorial Pacific Ocean during an El Niño. □

Ulrich Siegenthaler is in the Physics Institute, University of Berne, Sidlerstrasse 5, CH-3012, Berne, Switzerland.

1. Keeling, C. D. *et al.* in *Aspects of Climate Variability in the Pacific and the Western Americas* (ed. Peterson, D. H.) *Geophys. Monogr.* **55**, 165–236 (Am. Geophys. Union, Washington, D.C., 1989).
2. Volk, T. *Global biogeochem. Cycles* **3**, 267–279 (1989).
3. Siegenthaler, U. & Wenk, T. in *Extended Abstr. 3rd Int. CO₂ Conference Environmental Pollution Monitoring and Research Programme No. 59*, 189–194 (World Meteorological Organization, Geneva 1989).

Purifying earth

LAST week Daedalus proposed a novel deep-sea mining scheme, in which two boreholes were sunk in the sea floor to such a depth that the hot pressurized water would dissolve the rock between them. Thereafter, a convective flow would be established, with cold sea water sinking down one of the holes, and emerging from the other very hot, and saturated with minerals in solution. He now plans to try the same scheme on land.

DREADCO prospectors are studying hot-spring districts in Iceland and New Zealand, and the geothermal regions of southwest England, where hot rock is very near the surface. A borehole in such a region should soon reach a depth where water would boil but for the strong hydrostatic pressure above it. The hotter and more pressurized the water, the more ferocious it becomes as a solvent. Once the hole has reached a certain critical depth, it need only be kept full of water and it will dissolve its way deeper and deeper without further drilling. Two holes a suitable distance apart will ultimately join up by sideways diffusion of the water. A self-sustaining convective flow between them could then be easily established.

This simple arrangement would be a wonderful source of geothermal heat and deep-dissolved minerals. But Daedalus also sees it as a splendid new waste-disposal and sewage-treatment plant. Sewage, possibly loaded with suspended urban waste and industrial nasties like chlorinated residues, would simply be poured into the downflow pipe. In the heated and pressurized depths, all the solids would dissolve in the water and would begin to react with it and the surrounding rocks. Some of the metals in the waste, like sodium and aluminium, would probably be trapped by ion-exchange processes; others, like iron, copper, and manganese from spent batteries, would go to stable sulphide or carbonate minerals for reclamation. The organics, from paper and plastics to food waste and sewage solids, should hydrolytically oxidize, ultimately to carbon dioxide. All the waste of civilization would be disposed of simply and with total ecological rectitude!

The final outflow, with luck, would be a sterile solution of minerals in almost pure water. Once the minerals had precipitated out and the dissolved gases had fizzed away, it might almost be fit to drink! But even Daedalus is not quite so optimistic. He would prefer to arrange the pressure-gradient in the outflow so that the exit-water partly boils on its way up. The steam, after driving turbines for power, could be condensed to pure distilled water for domestic distribution. The residual mineral-loaded stream, concentrated by its partial evaporation, could be filtered to recover its metals.

David Jones

Crystals from first principles

SIR—Maddox¹ considers it a “continuing scandal” that “it remains in general impossible to predict the structure of even the simplest crystalline solids from a knowledge of their chemical composition”. In response, Cohen² has claimed that, on the contrary, “structures of crystals can be predicted using information about their chemical composition as the only input”. Without wishing to underestimate the significance of the results discussed by Cohen, I feel that he has greatly overstated the case.

The stated aim of first-principles calculations is to predict the structure and properties of a crystal given only its chemical composition. If calculations are started from a configuration close to the global energy minimum, this may be done very successfully for a variety of materials^{1,2}. Given the correct bond connectivity (or approximate relative atomic positions), cell size (in number of atoms) and (in some cases) space-group symmetry, the metric properties of the structure can be calculated, together with various properties of its energy and energy derivatives^{1,4}, but clearly these calculations require as input much more than simply the chemical formula. Although both Maddox and Cohen gloss over this problem, it is far from a trivial one.

There are two ways in which the calculations may be approached. First, an assemblage of randomly arranged atoms may be allowed to move such that the total energy of the system is minimized. This process can proceed with no assumptions (concerning, for example, which atoms are bonded to which, or what coordination numbers occur). The global energy minimum will be reached only if the derivative of the energy with respect to every variable parameter allows a continuous downward path towards it. From a random starting position, it is likely that the process will converge instead to a local minimum that represents a (possibly unphysical) metastable state. The calculations for silica³ discussed by Maddox show that several polymorphs are dynamically stable and will ‘trap’ any minimization that starts within their energy volume of influence. The actual occurrence of such metastable forms indicates that there is not a continuously decreasing energy path between the metastable and the stable forms. Simulated annealing techniques⁵ can permit escape from local minima, but are viable only if the number of local minima is small; this number is, however, usually very large. In general, therefore, solution is possible only by imposing strong boundary constraints, usually in terms of symmetry and cell size. For the former, a calculation starting with triclinic sym-

metry should converge on the correct symmetry. Imposing constraints on cell size is more difficult. Calculations can be performed for all possible cell sizes up to some arbitrary cutoff (with periodic boundary conditions), or for some large number of atoms (without periodic boundary conditions), where the expectation is that minimization will converge to an arrangement with the correct cell size and periodicity. In the latter case, the minimum number of atoms must be quite large because the properties of the assemblage must be those of a bulk crystal; experiments⁶ suggest that this requires something of the order of 100–1,000 atoms.

The alternative approach is to specify ‘good’ candidate structures to use as starting configurations, and this method is used most widely. It requires extensive assumptions about the material investigated, such as which types of atoms are bonded to which and what coordination numbers are present. Calculations have focused on materials that are quite well understood (silicon², silica^{1,3}, common minerals⁷) and the candidate structures used are the observed structures (under a variety of conditions) together with some other well-known arrangements of the correct stoichiometry. This selection of candidate structures can be considered neither rigorous nor exhaustive. Whether or not the minimization process incorporates directional chemical bonds, it is convenient to specify possible structural arrangements by their bond connectivity — the problem is then reduced to the study of periodic networks^{8,9}. Major problems include an insufficient understanding of fundamental constraints on the maximum cell size, or on the minimum and maximum allowable densities. For example, the observed distribution of 4-connected three-dimensional zeolite nets falls within a fairly restricted density range⁹. Even incorporating such restrictions, together with additional stereochemical information, the number of stoichiometrically possible networks can be very large. For example, the approximately 500 nets described by Smith⁸ could all be used as candidate structures for the silica polymorphs. In general, one needs realistic but more restrictive stereochemical constraints that reduce this number to more manageable proportions. Combinatorial studies of restricted structure types^{10,11} show that many potential structures occur even for simple stoichiometries. All of these are likely to correspond to local energy minima, heralding significant problems for an *a priori* energy minimization approach.

Thus a rigorous general solution to the question of crystal-structure prediction

may not be forthcoming in the near future. At least, for simple materials, it is possible to deduce accurate metric and physical properties from an approximate atomic arrangement, but much more work is needed in the area of predicting bond connectivities before the problem can even be considered as susceptible to solution.

FRANK C. HAWTHORNE

Department of Geological Sciences,
University of Manitoba,
Winnipeg, Manitoba,
Canada R3T 2N2

1. Maddox, J. *Nature* **335**, 201 (1988).
2. Cohen, M.L. *Nature* **338**, 291–292 (1989).
3. Tsuneyuki, S., Tsukada, M., Aoki, H. & Matsui, Y. *Phys. Rev. Lett.* **61**, 869–872 (1988).
4. Catti, M. *Phys. Chem. Miner.* **16**, 582–590 (1988).
5. Car, R. & Parrinello, M. *Phys. Rev. Lett.* **55**, 2471–2474 (1985).
6. Martin, T.P. *Angew. Chem.* **25**, 197–296 (1986).
7. Price, G.D., Parker, S.C. & Leslie, M. *Phys. Chem. Miner.* **15**, 181–190 (1987).
8. Smith, J.V. *Chem. Rev.* **88**, 149–182 (1988).
9. Brunner, G.O. & Meier, W.M. *Nature* **337**, 146–147 (1989).
10. Hawthorne, F.C. *Acta Crystallogr.* **A39**, 724–736 (1983).
11. McClarnen, T.J. *J. Solid State Chem.* **26**, 235–244 (1978).

Gene candidate

SIR—The subregional localization of malignant hyperthermia susceptibility (MHS) to chromosome 19q12–13.2 recently reported in *Nature*¹ can be used to suggest or eliminate ‘candidate’ genes as the site for the primary defect in this disorder. Indeed, MacLennan *et al.*² suggest, on the basis of linkage data, that a mutation in the ryanodine-receptor gene probably causes MHS. While awaiting the description of the mutation in MHS individuals which would confirm which gene is responsible for MHS, we would like to suggest another candidate.

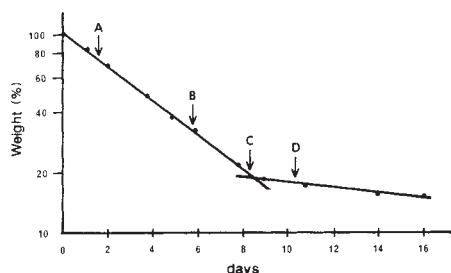
It is known that free fatty acid metabolism is abnormal in MHS skeletal muscle: muscle homogenates from humans³ and isolated mitochondria from pigs⁴ indicate that free fatty acid release is greater than two-and-a-half times that of normal controls. These free fatty acids seem to be derived from skeletal muscle triglycerides⁵. Elevated free fatty acid concentrations can increase the release and decrease the uptake of calcium in the sarcoplasmic reticulum⁵. The threshold for calcium-induced calcium release from sarcoplasmic reticulum is also reduced in MHS⁶, and fatty acids enhance the release of calcium by lowering this threshold⁷. Free fatty acids are also well known to produce heat by uncoupling oxidative metabolism.

A hormone-sensitive lipase is crucial in mobilizing free fatty acids from stored triglycerides and is regulated by hormonal and neuronal factors⁸. This lipase is activated by catecholamines through cyclic AMP-mediated phosphorylation, and dephosphorylated (inactivated) by insulin⁸. If this insulin inactivation did not occur in MHS individuals, excess free

1. McCarthy, T.V. *et al. Nature* **343**, 562–564 (1990).
2. McLennan, D.H. *et al. Nature* **343**, 559–561 (1990).
3. Fletcher, J.E. *et al. Br. J. Anaesth.* **58**, 1433–1439 (1986).
4. Cheah K.S. *et al. Biochem. biophys. Acta* **638**, 40–49 (1981).
5. Fletcher J.E. *et al. Eur. J. Anesth.* **6**, 355–362 (1989).
6. Nelson, T.E. *J. clin. Invest.* **72**, 862–870 (1983).
7. Fletcher, J.E. *et al. FASEB J.* **4**, A1209 (1990).
8. Holm, C. *et al. Science* **241**, 1503–1506 (1988).
9. Schonk, D. *et al. Cytogenet. Cell Genet.* **51**, 1075 (1989).
10. Schonk D. *et al. Genomics* **4**, 384–396 (1989).
11. Cytogenet. *Cell Genet.* **51(1–4)**, 1–1148 (1989).

1. Hubert, J.-C. *et al.* *EMBO J.* **2**, 2071–2073 (1983).
2. Hirashima, S., Hirai, H., Nakanishi, Y. & Natori, S. *J. biol. Chem.* **263**, 3858–3863 (1988).
3. Reinberg, D. & Roeder, R.G. *J. biol. Chem.* **262**, 3331–3337 (1987).
4. Sekimizu, K., Nakanishi, Y., Mizuno, D. & Natori, S. *Biochemistry* **18** 1582–1588 (1979).
5. Hirashima, S., Horikoshi, N., Sekimizu, K. & Natori, S. *Biochem. biophys. Res. Comm.* **160**, 1093–1099 (1989).
6. Rappaport, J., Reinberg, D., Zandomeni, R. & Weinmann, R. *J. biol. Chem.* **262**, 5227–5232 (1987).
7. Sluder, A.E., Greenleaf, A.L. & Price, D.H. *J. biol. Chem.* **264**, 8963–8969 (1989).
8. Horikoshi, M., Sekimizu, K. & Natori, S. *J. biol. Chem.* **259**, 608–611 (1984).

Though there have been many experiments concerning bound and free water reported¹⁻⁴, our method might reveal something new. There are known to be about 40 proteins in egg white; considering the molecular masses of the main ones⁵



Two stages of moisture evaporation in boiled egg white kept in an ordinary refrigerator and the state of its transparency. Ordinate, weight reduction of boiled egg white. The thin edges of boiled egg white became hard and transparent (vitrified) after the first day in a refrigerator, as shown by arrow A. After 6 days (arrow B) most of the boiled egg white had become hard and transparent. At arrow C (8 days) the evaporation entered the second stage. And finally, after 10 days (arrow D), all the egg white was hard and transparent.

and neglecting the rest we estimated the average molecular mass to be 51,000. The content of the free and bound water was estimated to be 1,900 and 19,000 moles per mole of protein, respectively, in the egg white. In other words, one gram of the protein contained 0.7 gram bound water and 7 gram free water. Assuming that the average molecular mass of amino acids is 110, it can be calculated that one amino acid contains 4 and 40 moles of the bound and free water, respectively.

The vitrified egg white was a plastic pale yellow solid. The plasticity could be due to the increased entanglement of the

denatured protein molecules. Vitrified raw egg white, on the other hand, is fine and forms thin fragments. X-ray diffraction of vitrified egg white reveals haloes characteristic of the amorphism seen in common glass, and the pore radius in the protein cluster was 6 angstroms. Neither the egg white nor fish eyeballs become putrified.

In the field of food sciences, it is very common to get dry protein product from its moistened material for preservation, as raw or processed products with high moisture are generally considered to putrify readily. As eggs are so common, it would be unusual and unnecessary to dry boiled eggs for preservation. Vitrification of denatured boiled egg white could therefore have been overlooked as a source of edible protein.

EISEI TAKUSHI

Department of Physics,
Faculty of Natural Science,

LIU ASATO

Faculty of Medicine and
Faculty of Agriculture,

TADASHI NAKADA

University of the Ryukyus,
Nishihara-cho,
Okinawa, 903-01 Japan

1. Pauling, L., *J. Am. chem. Soc.* **67**, 555 (1945).
2. Bather, R., Webb, S.J. & Cunningham, T.A. *Nature* **207**, 30 (1965).
3. Love, T.R. *Nature* **211**, 981 (1966).
4. Scheuplein, R.J. & Morgan, L.J. *Nature* **214**, 456 (1967).
5. Bahr, J.W. & Bakst, M.R., in *Reproduction in Farm Animals* (ed. Hafez, E.S.E.) 5th edn. 379 (Lea and Febiger, Philadelphia 1987).

Cautious searches

SIR—Skern *et al.*¹ are correct in stating that the results of computer searches for similarity between proteins must be interpreted with caution, but our paper² on the maize bifunctional inhibitor of α -amylase/trypsin merely suggested that both thaumatin and the tobacco pathogenesis-related protein should be re-examined for their possible effects on hydrolytic enzymes due to their strikingly extended sequence homology (52–57 per cent over 200 residues) with the inhibitor. Moreover, the sequence similarities we reported are of much greater statistical significance³ (Z values of 41.7 and 97.8) than those found by Skern *et al.* for the limited similarity between human rhinovirus 2 and thaumatin ($Z=5.8$) or trypsin ($Z=5.6$).

The validity of our suggestion has been supported by several recent discoveries. For example, a potent inhibitor of locust gut α -amylase isolated from seeds of *Coix lachryma-jobi* is, in fact, also an endochitinase⁴. We were prompted to look for the enzyme activity only when a computer search revealed that the inhibitor had a similar sequence to endochitinases. Simi-

larly, when a putative α -amylase/protease inhibitor isolated from barley seeds^{5,6} was found to have sequence similarities with phospholipid transfer proteins⁷, subsequent assays confirmed that it possessed this function⁸. The further observation that an amylase inhibitor from beans is similar to lectin (agglutinin)-like proteins⁹ tends to confirm the belief that computer-aided searches will continue to provide a valuable indication of the possible, but as yet unsuspected or unproven, multifunctional nature of many defensive plant proteins.

M. RICHARDSON

Department of Biological
Sciences,
University of Durham,
South Road,
Durham DH1 3LE, UK

1. Skern, T., Zorn, M., Blaas, D., Kuechler, E. & Sommergruber, W. *Nature* **344**, 26 (1990).
2. Richardson, M., Valdes-Rodriguez, S. & Blanco-Labra, A. *Nature* **327**, 432–434 (1987).
3. Lipmann, D.J. & Pearson, W.R. *Science* **227**, 1435–1441 (1985).
4. Ary, M.B., Richardson, M. & Shewry, P.R. *Biochim. biophys. Acta* **993**, 260–266 (1989).
5. Svensson, B. *et al. Carlsberg Res. Commun.* **51**, 493–500 (1986).
6. Mundy, J. & Rogers, J.C. *Planta* **169**, 51–63 (1986).
7. Bernhard, W.R. & Somerville, C.R. *Archs biochem. Biophys.* **269**, 695–697 (1989).
8. Breu, V. *et al. Calsberg Res. Commun.* **54**, 81–84 (1989).
9. Moreno, J. & Chrispeels, M.J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7885–7889 (1989).

Killer toxins

SIR—White *et al.*¹ have reported that the yeast killer toxin encoded by the *Kluyveromyces lactis* plasmid pGKL1 is not an inhibitor of adenylyl cyclase, leaving open the question of the mechanism of irreversible growth inhibition in sensitive yeast cells. Recent experience with endochitinase genes from plants and exochitinase genes from bacteria may shed some light on the mode of action of the yeast killer toxin.

Many plants have genes that encode enzymes with chitinase activity; I am most familiar with chitinase-related genes from hybrid poplars². A TFASTA (ref. 3) sequence-similarity search of the GenBank 59.0 DNA (ref. 4) and PIR 21.0 protein (ref. 5) sequence databases using a chitinase-related (Win8) protein sequence (deduced from its complementary DNA sequence) revealed the previously reported global similarity to other plant chitinases and local similarity to plant chitin-binding proteins².

Unexpectedly, the large subunit of the yeast killer toxin from pGKL1 (ref. 6) shares a region corresponding to the chitin-binding domain with this collection of plant proteins. When I used the toxin protein itself to search the sequence databases, an additional similarity (outside the domain shared with plant chitin-binding proteins) to the exochitinase from the bacterium *Serratia marcescens* (ref. 7) uncovered was that *Serratia* chitinase has no discernable sequence identity with plant chitinases.

Because fungal cell walls are made predominantly of chitin, and plant and bacterial chitinases are effective antifungal agents^{8–10}, it seems reasonable to suggest that the large subunit of the *K. lactis* killer toxin arrests the growth of target yeast cells by interfering with cell-wall biosynthesis. It may be useful to test the purified toxin both for the predicted chitin-binding and chitinase activities. If the toxin has exochitinase activity, the domain in common with the chitinase from *Serratia* could be the site of glycosidic bond cleavage.

HARVEY D. BRADSHAW, JR

Department of Biochemistry,
University of Washington,
Seattle, Washington 98195, USA

1. White, J. H., Butler, A. R. & Stark, M. J. R. *Nature* **341**, 666–668 (1989).
2. Parsons, T. J., Bradshaw, H. D. Jr & Gordon, M. P. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7895–7899 (1989).
3. Pearson, W. R. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2444–2448 (1988).
4. Bilofsky, H. S. *et al. Nucleic Acids Res.* **14**, 1–4 (1986).
5. George, D. G., Barker, W. C. & Hunt, L. T. *Nucleic Acids Res.* **14**, 11–15 (1986).
6. Stark, M. J. R. *et al. Nucleic Acids Res.* **12**, 6011–6030 (1984).
7. Jones, J. D. G. *et al. EMBO J.* **5**, 467–473 (1986).
8. Schlumbaum, A. *et al. Nature* **324**, 365–367 (1986).
9. Mauch, F., Mauch-Mani, B. & Boller, T. *Plant Physiol.* **88**, 936–942 (1988).
10. Roberts, W. K. & Selitrennikoff, C. P. *J. gen. Microbiol.* **134**, 169–176 (1988).

Sensible mathematics

SIR—Dirac¹, after describing the electron self-energy problem and mass renormalization, states, "Most physicists say that QED is a good theory. I must say I am very dissatisfied with the situation because this so-called 'good theory' does involve neglecting infinities which appear in its equations, neglecting them in an arbitrary way. This is just not sensible mathematics. Sensible mathematics involves neglecting a quantity when it is small. . . not just because it is infinite and you do not want it."

In agreement with Dirac, we believe that perturbation theory must be altered or reconstructed into such a form (without spoiling relativistic invariance) that the self-energy integral (in momentum space) taken over the momenta of all virtual photons converges. Feynman, in his paper on QED, multiplied the photon propagator, k^{-2} , by the *ad hoc* factor $-f^2/(k^2 - f^2)$, where k is the momentum of the virtual photon (essentially its frequency) and f is an arbitrarily large parameter. Although this convergence factor, which is essentially equivalent to a high-frequency cutoff, preserves relativistic invariance, it is objectionable because of its *ad hoc* character and the presence of f without any theoretical justification. We show below, however, that the Feynman convergence factor appears in the self-energy integral naturally if general relativity is taken into account.

Isham *et al.*² have shown that in a 'gravity-modified' QED the "inverse of the gravitational constant appears as an effective cut-off mass" in the calculation of the electron self-energy, but their calculation is so encumbered by formalism that it is difficult to see the physics that is involved. For that reason, our calculation below, which leads essentially to the same result and to a general relativistic Feynman convergence factor, is important. Our calculation takes into account the gravitational reddening suffered by a virtual photon owing to the increased mass of the emitting particle produced by its recoil momentum.

Because momentum, unlike energy, is conserved in each single virtual process (each integral for a single virtual photon is taken over all configuration space, so that there is no uncertainty in momentum, as emphasized by Dirac¹) the frequency ν_0 of the virtual photon, just when it is emitted, is given by the equation $2\pi\hbar\nu_0/c = m_0/v[1 - (\nu/c)^2]^{1/2}$; here m_0 is the resting mass of the emitting particle, v is its recoil velocity, and $m = m_0/[1 - (\nu/c)^2]^{1/2}$ is its effective mass, which redshifts the frequency of the virtual photon in accordance with the Einstein formula $\nu = \nu_0[1 - (2Gm/c^2r)]^{1/2}$, where r is the Compton wavelength $\hbar/mc$

of the emitting particle. From our first equation (the conservation of momentum), we have for the recoil velocity

$$\nu = (h\nu_0/m_0c)[1 + (h\nu_0/m_0c^2)^2]^{1/2}$$

and

$$m = m_0[1 + (2\pi\hbar\nu_0/m_0c^2)^2]^{1/2}$$

On substituting this into the equation for the effective frequency of the virtual photon, we obtain

$$\nu = \nu_0[1 - \frac{2Gm_0^2}{\hbar c} - 8\pi^2 G\hbar \nu_0^2/c^5]^{1/2}$$

Because the second term $2Gm_0^2/\hbar c$ in the parenthesis is negligible for the electron, we can exclude it and write $\nu = (f^2)^{-1/2}(f^2 - \nu_0^2)^{1/2}\nu_0$, where $f^2 = c^5/8\pi^2 G\hbar$, a universal constant (the square of a frequency), is of

the order of 10^{95} and therefore extremely large, as required by Feynman. Multiplying f by h and dividing by c^2 , we obtain the universal mass $m = (\hbar c/G)^{1/2}$ as the cut-off mass. This is just the Planck mass, which indicates that gravity is important in perturbation theory.

LLOYD MOTZ
ROBIN O. MOTZ

Rutherford Observatory,
College of Physicians and Surgeons,
Columbia University, New York,
New York 10027, USA

1. Dirac, P. A. M. *Directions in Physics* (Wiley, Chichester, 1978).
2. Isham, C. J., Salam, A. & Strathdee, J. *Phys. Rev. D* **3**, 1805 (1971).

Short alpha-helix stability

SIR—Because C-peptides, the 13-residue peptides simulating an α -helix of RNase, are significantly helical in aqueous solution near 0 °C^{1,2}, the helicity has been attributed to specific side-chain interaction. The helicity is 30 times more than that predicted by Zimm-Bragg theory with host-guest stability parameters³.

Yet the observed helicity of C-peptides (between 15 per cent and 45 per cent depending on amino-acid composition and pH) is close to that expected for uncharged polylysine fragments of the same length, where specific side-chain interactions (salt bridges, hydrogen-bonding, aromatic interactions) are absent.

By direct measurement⁴, the helix-stability parameters for uncharged polylysine have been found to be $s_a = 1.3$ (at

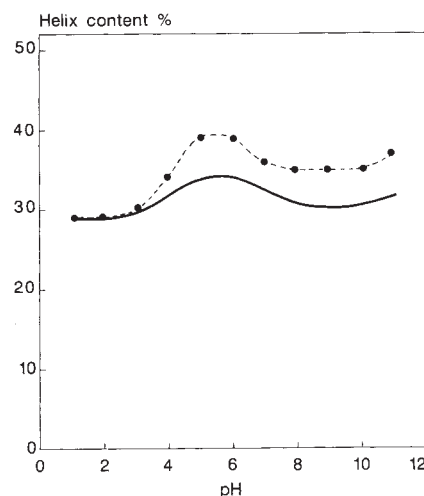
about 0 °C) and $\sigma_a = 0.0025$. Using these values, the Zimm-Bragg theory⁵ gives 30 per cent helicity for 13-residue polylysine with CONH groups at either end (as in C-peptides). Thus the 'general level' of the observed helicity of C-peptides is normal (at about 0 °C) for a short peptide consisting of helix-forming residues without the intervention of specific side-chain interactions.

At room temperature, the helicity of polylysine fragments must be half as much; s_a falls⁴ to 1.14 at 25 °C, similar to the decrease of helicity found for the C-peptides¹. It is relevant that helicity of about 5 per cent has been predicted⁶ for the corresponding fragment of unfolded RNase.

The helicity of a peptide differs from the 'general level' because of a multitude of specific side-chain interactions, (usually rather weak), many of which (usually very weak) have been taken account of theoretically^{6,7}. Thus the 1.2-fold increase of helicity of C-peptides on ionization (see figure) is explained by the electrostatic attraction of Glu⁻ and His⁺ to the partial charges at the nearest termini of a fluctuating α -helix. Present theory^{6,7} may thus serve as a guide for singling out the more specific effects observed in experiments with short peptides².

ALEXEY V. FINKELSTEIN
AZAT YU. BADRETDINOV
OLEG B. PTITSYN

Institute of Protein Research,
USSR Academy of Sciences,
Pushchino,
Moscow Region,
USSR



pH dependence of α -helical content for the C-peptide (acetyl-AETAAKALRAHA-amide) at 3 °C in 0.1 M NaCl. Solid line, experiment⁴; dotted line, theory⁷. The temperature change of s_a was calculated using the helix-coil transition enthalpy $\Delta H = -1.1$ kcal mol⁻¹ (see ref. 4) for all amino acids. All other parameters are independent of temperature^{6,7}.

1. Shoemaker, K.R., Kim, P.S., York, E.Y., Stewart, J.M. & Baldwin, R.L. *Nature* **326**, 563-567 (1987).
2. Shoemaker, K. R. *et al.* *Cold Spring Harbor Symp. Quant. Biol.* **LII**, 391-398 (1987).
3. Sueki, M. *et al.* *Macromolecules* **17**, 148-153 (1984).
4. Barkaya, T.V. & Ptitsyn, O. B. *Biopolymers* **10**, 2181-2197 (1971).
5. Flory, P.J. *Statistical Mechanics of Chain Molecules* chapter 7.5 (Wiley, New York, 1969).
6. Finkelstein, A.V. & Ptitsyn, O. B. *J. molec. Biol.* **103** 15-24 (1976).
7. Finkelstein, A.V. *Program ALB Brookhaven Protein Data Bank and EMBL* (1983).

Translation from the Greek

V. Reynolds

Oedipus in Evolution. By Christopher Badcock. *Basil Blackwell: 1990. Pp. 221. £29.95, \$37.95.*

I ASSUME most readers of *Nature* feel they can get on quite well without Freud. Perhaps, indeed, most people can. Yet we are told incessantly by media pundits and even some academics and "talking heads" that our whole way of looking at ourselves this century is Freudian, that Freud has changed us. So the question arises whether modern science has anything to say about this, whether we must just accept that we are somehow post-Freudian, or whether we can do better than that.

Badcock's self-appointed task, started in earlier volumes and continued in this one, is to show that Freud was right, and to show this by recourse to sociobiology. His technique is to demonstrate that things like the Oedipus complex, penis envy, castration complex or regression, are Freud's way of describing what sociobiologists like Robert Trivers have subsequently shown to be evolved propensities, the outcome of natural selection.

Who was Oedipus? A character in a Sophoclean play who unwittingly killed the King his father, as a result of a prophecy, where three roads met, on the way from Thebes to Delphi. What was Oedipus thinking of, to kill his father? He wasn't thinking, he didn't even know; he was driven, driven by fate, by prophecy. Link number one with evolution; to be driven, not to be a wholly free agent; evolution has two drivers, survival and reproductive success.

Oedipus later found out that it was indeed his father he had killed and, worse, that he had gone on to marry his own mother. So deep was his shame that he put out his eyes. In addition, he had brought famine and despair to the people of Thebes, and blinding himself expiated his sin. This is Greek tragedy at its finest. Badcock's preoccupation with sociobiological reinterpretation must pale by comparison, and indeed so must Freud's own work, though that at least has elements of the bizarre about it.

Freud took the Sophoclean story and made of it an interpretation of an aspect of childish behaviour, the longing of the young boy to possess his mother exclusively and to oust his father from the family. Each of us, said Freud, and Badcock accepts this without question, goes through this phase, after the earlier oral one. Having emerged from a total preoccupation with sucking the mother's nipple, boys start to get sexually possessive, and mothers respond warmly to their

offspring's amorous advances. Why should these things (shocking to Freud's generation, boring to ours) go on? Because, says Badcock, young children have been selected during evolution to behave in ways that deter the mother from

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

The infant Oedipus rescued from the mountain by a shepherd, where Apollo had warned that Oedipus would kill his father.

investing in another child, so as to maintain the maximum amount of parental investment for themselves.

In Freud's equation, Oedipus is you and me; we are driven to (want to) kill our father and to (want to) marry (that is, possess) our mother. The twin pillars of the equation are: unwitting = unconscious and driven = evolved. The unconscious human psyche has evolved in such a way that now, during childhood, as it develops, it sets itself these two unattainable goals — patricide and incest. Does sociobiology really support this? According to Badcock it does; Freud has been vindicated by sociobiology.

I am arguing that Oedipal behaviour in little boys . . . should be looked at from the point of view of its costs and benefits to the individual child in terms of parental investment. According to this way of looking at things, Oedipal behaviour would constitute a kind of "rehearsal" or "advertisement" of the boy's potential reproductive success, aimed at securing preferential parental investment because of the potentially much greater reproductive success which an individual male can have in the context of a polygamous mating system. (Page 87.)

(Polygamy is here used to include officially monogamous systems in which there is extramarital sex.) We know the

The Mansell Collection
sociobiological arguments by now, and it is rather tedious to have them explained all over again. I would not want to question whether sociobiology itself has anything useful to tell us about the human condition, or the processes of child development; I take it as given that it has. We can readily accept that young children behave in ways that maximize their access to parentally derived benefits, that is, the facts of

sibling rivalry and parent-offspring conflict as applied to ourselves.

But can we accept that this is what Freud was saying all along? I don't see how we possibly can. Sociobiology has its roots in biology, and therein lies its strength. It derives its concepts and theories from the ideas of Darwin and those who have followed after. Freud, writing about Oedipus in 1900, might have, could have followed Darwin, but he chose a very different path — an intuitive path not a scientific one — based on his own idiosyncratic interpretations of people's dreams and

neuroses. Freud thus has to be classed with those thinkers who eschewed the scientific tradition, not with those who embraced it. Today, Freud is history. If some of his insights seem akin to the findings of sociobiology, he is not thereby vindicated, other than in the sense that one or two of his hunches were near the mark. Badcock demonstrates commendable loyalty to the old man, but I fear that there is nothing anyone can do to put him back on his pedestal. □

V. Reynolds is in the Department of Biological Anthropology, University of Oxford, 58 Banbury Road, Oxford OX2 6QS, UK.

State of the art

David Bailin

Electroweak Interactions. An Introduction to the Physics of Quarks and Leptons.

By Peter Renton. Cambridge University Press: 1990. Pp. 596. £60; \$110.

QUANTUM electrodynamics (QED) is a relativistic quantum field theory formulated by Feynman, Dyson and Schwinger around forty years ago. Its incredible success in describing and predicting purely electromagnetic processes with apparently limitless accuracy in terms of very few parameters, naturally suggested that the other interactions of particle physics, the weak and strong interactions, should also be described by relativistic quantum field theories. We now know that this expectation is correct, and that the weak and electromagnetic interactions are described by a unified gauge field theory called electroweak theory, while the strong interactions are described by quantum chromodynamics (QCD).

Electroweak theory differs from QED in several important respects. First, the gauge symmetry group ($SU(2) \times U(1)$) underlying the theory is larger than the simple $U(1)$ group (of transformations on a circle) which underlies QED. Second, this group is non-abelian, meaning that it is non-commutative. Finally, the theory is 'chiral', reflecting the fact that weak interactions, unlike electromagnetic ones, are not invariant under the parity transformation: we can say that left and right chiral states have different weak charges. None of these differences is much more than technical, and in principle an $SU(2) \times U(1)$ invariant gauge theory could have been formulated at any time after about 1954. The really vital difference from QED is that the symmetry is broken. If it were not, the weak gauge particles, the W and Z bosons, would be massless, just as the photon is and, because the theory is chiral, the electron, muon and other fermions would also be massless. Breaking the gauge symmetry is even easier than preserving it, but in general the breaking destroys the 'renormalizability' of the symmetric theory. In a renormalizable theory, such as QED, the theory is rendered finite by absorbing (ultraviolet) divergences into a finite number of (unpredicted) parameters. The problem with breaking the symmetry by hand, as it were, is that an infinite number of such renormalizations is needed, and the theory loses the ability to predict, which is the hallmark of its paradigm (QED). The trick, we now know, is to break the symmetry 'spontaneously'; the dynamics preserves the symmetry, but the ground state (the particle-physics vacuum) breaks it. (It was Weinberg's use of this

technique on Glashow's gauge symmetric lagrangian which finally led to the electroweak quantum field theory, whose renormalizability was demonstrated in the early 1970s.)

Most of the components of this theory have now been tested against experiment with overwhelming success. It is the renormalizability of the theory which permitted the prediction of the W and Z boson masses to within about $1 \text{ GeV } c^{-2}$ of the values measured when the particles were triumphantly discovered at CERN in 1983. Similarly, the theory now provides upper and lower limits on the top quark mass before its discovery. The final 'proof' that the symmetry is spontaneously broken must await the discovery of the Higgs boson, whose mass is unfortunately not predicted by the theory. The undeniable success of the theory naturally requires that graduate students of particle physics learn the theory and how it works. This demands not merely the study of quantum field theory, and QED in particular, but also the novel features of electroweak theory to which I have alluded. Increasingly it falls to experimentalists to perform the theoretical calculations against which the data are tested, and which 15 years ago were the preserve of theoreticians.

It is therefore entirely appropriate that this book on electroweak theory is by an experimentalist, in fact one who is currently engaged in precision tests of the theory at the DELPHI detector on LEP, the electron-positron collider opened at CERN six months ago. Having an experimentalist as its author, the book is naturally replete with an abundance of very impressive data detailing just how successful the model has been. The problem for the author of such a book is where to begin, and Renton goes right to the very beginning, dealing with wave equations, field operators and QED, including its renormalization, in the first third of the book. All these topics are of course well covered (in more detail) in many books, some more than 30 years old. In petroleum retailing it is well known that if a crossroads (say) has three filling stations, then this demonstrates that the location is ideal for selling petrol and that a fourth filling station will also thrive. As it is in gas stations, so it is apparently in bookselling. Renton declares in his preface that "the main emphasis is on the understanding of physical processes in terms of lowest order calculations". He stresses that higher order calculations are important and refers the reader to "more advanced texts" for details. I think this is a pity. The renormalization is rather complicated, as the novel features of electroweak theory require techniques which are not required in QED. Nevertheless, as the whole point of the theory is that it is renormalizable, and therefore

admits precision tests, and since experimentalists are beginning to do these calculations, I think that some attempt at covering this topic should have been made, probably dispensing with the earlier material. Another theorist's gripe would be the failure to cover the 'chiral anomalies'. It is precisely the requirement of 'anomaly cancellation' which makes us so certain that the top quark must exist. Not many books discuss either of these, so perhaps my (academic's) criticism can be rebutted by the (retailer's) response that the market is not yet ready for such an innovation.

In any case, my overall judgement is very positive, and I think that the book will be well received, especially by experimentalists. It is about much more than electroweak interactions. (It includes QCD, and the possibilities beyond the standard model.) In many ways the subtitle conveys more accurately than the main title what is in the book: an (excellent) account of the current state of particle physics. □

David Bailin is in the School of Mathematical and Physical Sciences, University of Sussex, Brighton BN1 9QH, UK.

Culture club

R. H. A. Coutts

Genetic Engineering of Crop Plants.

Edited by G. W. Lycett and D. Grierson. Butterworths: 1990. Pp. 293. £65.

IN the first chapter of this collection of papers presented at the 49th Nottingham Easter School in Agricultural Sciences, Edward Cocking illustrates the importance and the impact of plant cell and tissue culture on plant genetic engineering; most of the subsequent contributions serve only to reinforce this statement. In some cases, where researchers are dealing with plants amenable to tissue culture and which can be regenerated, such as tobacco and potato, the results are exciting. But with other, more recalcitrant species, such as some monocotyledonous plants and legumes, optimism abounds.

The scope of the book is certainly wide ranging, from coat protein-mediated protection against plant virus infection to nodulation genes of *Rhizobium*. In fact these chapters, respectively from the laboratories of Roger Beachy and Tony Johnston, are highpoints in the book: the assembly of useful data includes not only that obtained by the named researchers efforts but reviews that of others. The underlying enthusiasm of molecular biologists, especially through the discovery of transgenic plant technology, is evident in the crisp and stylish presentation. ►

What sets this book aside from the usual collection of conference proceedings is that the contributors not only present their own recent data but have tried to integrate their thoughts on other research in their sphere of interest. Such presentations were particularly evident in the contribution from the Durham plant molecular biology group.

I am unsure as to whether the absence of methodology sections from the chapters will appeal to the purists, but good reference lists at the end of most chapters should encourage the curious. The book is generally well-illustrated with good pictorial evidence of results, particularly in the presentations by Stanford *et al.*, on wound-induced genes in potato, and Covey *et al.*, on pathogenesis of cauliflower mosaic virus.

One undeniable problem of publishing the proceedings of any scientific meeting is the inevitable time difference between presentation and publication. Nevertheless, the editors have assembled a lucid and informative volume within a year, which is to be applauded. Even so, some of the contributions have or are about to appear in other volumes, which detracts a little from this one. Also, the competitive nature of the science may have deterred some, but not all, of the participants from divulging new information — again, an unavoidable problem. To those in the business of teaching the genetic engineering of crop plants, the synthesis of complex data and a volume of research papers into succinct accounts of exciting science will be extremely useful.

It has become *de rigueur* to publish the proceedings of international conferences almost as an afterthought. Because of this fact many researchers' and lecturers' shelves are weighed down by an ever increasing number of books which rapidly lose their scientific value, either because they are out of date and do not reflect current scientific dogma or they are badly written and poorly presented. It is a pleasant change to read a book which does not have these pitfalls, and I hope that it is well read and well used by students and researchers alike. Indeed, I feel it is these areas of higher education, especially at the undergraduate and first-year post-graduate stages, where the book will find its most enthusiastic reception. □

R. H. A. Coutts is in the Department of Biology, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BB, UK.

Correction

Information supplied by the Science Photo Library for the picture accompanying the review of *Science-based Dating in Archaeology* (*Nature* **345**, 126; 10 May 1990) incorrectly stated that the Oxford radiocarbon accelerator could date samples up to 100,000 years in age. This should have read 40,000 years. □

Class act

J. Z. Young

Cephalopodes. Edited by Katharina Mangold. *Traité de Zoologie*. P-P. Grassé. Tome V, Fasc. 4. Masson: 1989. Pp. 804. FF1,100. (In French.)

OCTOPUSES, squids and cuttlefishes are among the most interesting inhabitants of the seas, both scientifically and economically. Moreover, the relics of their ancestors in the rocks are the most familiar of all fossils. This excellent book about them is the latest edition to Grassé's monumental treatise, now extending to 39 tomes and covering the whole animal kingdom.

Anyone writing a treatise on Cephalopods has to please two distinct populations of readers — zoologists and palaeontologists. These groups have different aims, use different languages and even have different views of scientific method. Yet the work of each group requires knowledge of the results of the other. The volume under review is really two books, a long one by zoologists, led by K. Mangold, and a shorter, by the palaeontologist C. Teichert. It was first undertaken by A. Portmann and has been continued by Mangold, who has enlisted the help of others but overseen the whole herself, so that happily it is not a multi-author book.

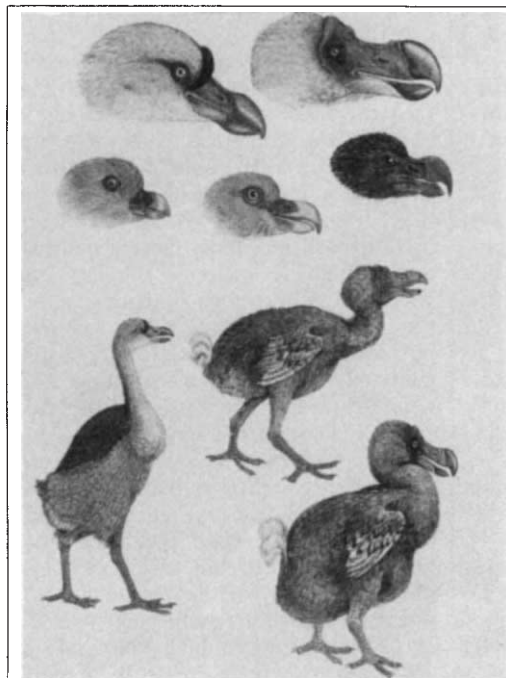
With her great knowledge of cephalopods, especially those of the Mediterranean, and of the literature, Mangold has produced a very satisfactory survey of the living squids, cuttlefishes and octopuses. The book is beautifully produced, with many excellent illustrations. Line drawings have been ingeniously copied from half-tones, though sometimes with distor-

tion of the cell pattern; for instance in nervous tissue. There are full bibliographies and a useful index.

This book has been in preparation for many years. The zoological part has been kept up to date, carrying references up to 1987. The palaeontology section, however, goes only to 1983. But the point of such a work is not that it includes the latest news or fresh original data, but that it provides a thorough view of the many aspects of the animals. Besides description of the structure and functioning of the various organ systems and their embryology, there are valuable ethological chapters on distribution, migrations, parasites and predators, including fishermen (but curiously little about prey). The systematic section briefly reviews every genus and will be valuable in itself.

Cephalopods provide the material for an increasing volume of biological research, ranging from molecular biology to neurophysiology. Workers who are studying, say, the pigments of blood or the membranes of giant nerve fibres can learn in this book the place of their tissues in the life of the octopus or squid. They may not find much detail about their own topic: for instance a footnote warns that works on the giant nerve fibres and synapses are "*extrêmement nombreux et importants, mais ils dépassent le cadre d'un Traité de Zoologie*".

The section on the fossil Nautiloids and Ammonoids compresses a huge controversial subject into 70 pages. It will probably be too summary for palaeontologists and is rather hard going for zoologists. The author discusses the origin of the whole class and of the various groups, but despite the vast number of fossils there is much controversy and few conclusions about the causes of their evolution and



These drawings of the dodo *Didus ineptus* (top, middle and lower right), the Solitaire *Pezophaps solitaria* (top and bottom left), and the mainland Green Pigeon, Sao Tome Green Pigeon and Tooth-billed Dove (three heads in the middle), are by Jonathan Kingdon, taken from his recent book *Island Africa*. Full of beautiful illustrations by the author, the book documents the evolution of Africa's rare animals and plants. It is a record of the treasures of species to be found in enclaves throughout the continent. Published by Collins, price £25.

extinctions. The function of the shell is largely to regulate buoyancy and it is a pity that more has not been deduced from its various shapes about the mode of life of the animals that lived in them, and the reasons for their success or failure.

Teichert discusses the eight 'crises' in which the ammonoids nearly became extinct before they finally disappeared at the end of the Cretaceous. The crises were very irregular, occurring at intervals varying from 5–120 million years, curiously not affecting the nautiloids at the same times. The author considers and rejects various hypotheses put forward to explain this, considering that multiple causes were at work and that most authors have been "trop simpliste" in their discussions.

One of the uses of a treatise such as this

is to provide classifications, and it is unfortunate that those given in the two parts of the book do not agree. Nor is there any clear view of the evolutionary relationship of the various groups. It is particularly disappointing that in neither part is there any discussion of the relationship of the various orders or families of modern cephalopods — admittedly a difficult subject. The difficulty partly results from the shocking scarcity of data about genetics within the group. The authors are not to blame for this. They have produced a splendid book, which will be useful for many years. □

J. Z. Young is in the Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX3 0DN.

Radical thought

Asa Briggs

The Invention of Progress. The Victorians and the Past. By Peter J. Bowler. *Basil Blackwell: 1989. Pp 231. £25, \$34.95.*

MUCH, though not all, of the voluminous writing on "the Darwinian revolution" fails to do justice to other nineteenth-century intellectual developments that enable the revolution to be placed in perspective. In linking developments in biology to developments in geology, anthropology, archaeology and, not least, history and pre-history, Peter Bowler is not perhaps as original as he claims. Yet he has written a comprehensive, stimulating and ambitious book which starts not with biology but with history and which moves on via "the fascination of the past" and through sociology, philology, theology (less fully treated, though often touched upon) and law before turning to "the antiquity of man", "missing links", "the origin of races", "fossils and progress", "the ascent of life", "the record of the rocks" and "progress by leaps". The attractive title "*The Invention of Progress*", backed by well-chosen illustrations as well as text, points to the author's unifying explanation of what was written and said in a remarkably impressive range of often heated Victorian debates. "The idea of progress was imposed upon history to create the sense of order that the Victorians craved. In a very real sense they *invented* the past, since all the factual discoveries of antiquaries and archaeologists had to be interpreted within a conceptual scheme that satisfied the cultural demands of the age." "Only by looking across the whole field of Victorian historical studies", Bowler argues, can the content and implications of the different but related Victorian debates be fully understood.

Bowler's account of many of the particular debates makes fascinating reading even when the reader is familiar with the content of the debate in question. He is a clear expositor and knows well how to summarize different points of view and how to chart the way in which the points of view of particular individuals changed. Yet it is easier for readers asked for the first time to consider the whole range of the debates to take some of the arguments more seriously than others. The author's approach is that of an intellectual, not a social historian, and he does not always tell us how seriously or how widely some of the debates influenced contemporaries. Moreover, he does not discuss how far in a period of increasing, if still at the end of the period imperfect, academic specialization there were people who were interested in all the topics focused upon. Is he assembling material that was not assembled at the time? His own sense of purpose in trying to relate everything to everything else is at least as determined and as powerful as that sense of "cosmic teleology" which he imputes to most of the writers with whom he is concerned.

In treating progress as the key to everything, Bowler is at pains to insist that there was no unanimity about its character, and in relating progress to history and science he is himself inevitably leaving out economics and technology, subjects that generated other kinds of debate about the nature of progress. For his own purposes he finds it necessary to distinguish only between two nineteenth-century approaches to "progress" — the first sequential or continuous; the second episodic or cyclical. The latter, according to him, for long predominated. Thus, the belief that each "race" — and he is especially illuminating on this difficult subject — had contributed to "the ascent of man", only to give way before "higher types", was paralleled in the writings of "palaeologists opposed to gradual evolution". In each case, Bowler seems to be arguing,

the reasons why the episodic or cyclic view was for a long period the dominant view lay outside intellectual history. The view was significant because it made it possible for Victorians either to continue to believe in "a divine plan" or in "imperial mission". "Like the great empires of the past, each of the vertebrate classes had had its period of glory before being displaced by the next highest manifestation of the Creator's power."

In the last few densely packed pages of the book, which deal with the decline of the idea of progress in the late-nineteenth century and twentieth-century — "degeneration", cultural relativism, the "new psychology", "Spenglerism" — Bowler argues also that it was factors outside intellectual history as much as changes in the content of the sciences that accounted for the change. "It is difficult to believe that the belated recognition of the full potential of Darwin's theory in the early-twentieth century played more than a minor role in the crisis of confidence that eventually destroyed the Victorian concept of progress."

It should be added that *The Invention of Progress* incorporates a thesis which the author has already advanced in an earlier study, *The Non-Darwinian Revolution*, published in 1988, a thesis with two conclusions. First, while *The Origin of Species* played a major "catalytic role in converting the world to evolutionism", late-nineteenth century evolutionism remained "emphatically developmental in character" and "natural selection remained peripheral to most biologists' views on the mechanism of evolution until well into the twentieth century". "The open-ended undirected character of evolution in Darwin's original theory was ignored or misrepresented". Second, before Darwin there was "an active movement by political radicals and younger medical men to exploit the materialist implications of evolutionism in their campaign against the political, academic and medical establishments, and Darwin himself was "unable or unwilling" to "sever the link between evolutionism and the developmental way of thinking". "Radical materialism" was, in consequence, held back.

Both parts of the thesis are stimulating. The second, however, raises questions of the individual motivation of scholars — and others — which require further discussion. "Victorian values", which figure in Bowler's index, themselves remain matters of debate. So, too, does "materialism", "radical" or otherwise. □
Asa Briggs is Provost of Worcester College, Oxford OX1 2HB, UK.

■ Also recently published, *In Pursuit of a Scientific Culture* by P. A. Dale, documents the search of Victorian writers for a philosophical replacement for romanticism. Published by University of Wisconsin Press. Price, hbk \$42.50; pbk \$17.50.

Evolution and detectability of comet clouds during post-main-sequence stellar evolution

S. Alan Stern^{*†}, J. Michael Shull^{†‡} & John C. Brandt^{*†}

^{*} Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, Colorado 80309-0392, USA

[†] Department of Astrophysical, Planetary, and Atmospheric Sciences, University of Colorado, Boulder, Colorado 80309-0391, USA

[‡] Joint Institute for Laboratory Astrophysics, University of Colorado and National Institute of Standards and Technology, Boulder, Colorado 80309-0440, USA

Stars leaving the main sequence enter a luminous phase, during which circumstellar comet disks would be partially evaporated. Such a process could explain some observations of water masers and complex molecules in envelopes of giant and supergiant stars, and might provide a means of searching for Oort-type clouds—and by inference, past planetary formation—around other stars.

IN some $4\text{--}5 \times 10^9$ years the Sun will exhaust its hydrogen supply and begin its post-main-sequence evolution, a common process for lower main-sequence stars¹. During post-main-sequence evolution, late-type stars typically attain luminosities between $6 \times 10^3 L_\odot$ and $6 \times 10^4 L_\odot$ for periods of one million to several tens of million years. We develop a time-dependent model for the destruction of comets within the sublimation radius of typical post-main-sequence stars, and show that the effects of such a luminosity increase on the star's reservoir of comets will be dramatic. We also point out that comet-cloud destruction produces several detectable signatures. If comet clouds are common around solar-type stars, then prodigious cometary destruction should be an observable aspect of post-main-sequence stellar evolution.

We begin by evaluating the radiative equilibrium temperature T_e of a unit surface normal to the post-main-sequence insolation at astrometric distance R

$$T_e(R) = \left[\frac{(1-A)L_{\text{poms}}}{4\pi\epsilon\sigma R^2} \right]^{1/4} \quad (1)$$

where σ is the Stefan-Boltzmann constant, A is the surface bolometric Bond albedo, ϵ is the comet surface emissivity, and L_{poms} is the luminosity of the post-main-sequence star. Using representative parameters of $A=0.04$, $\epsilon=0.9$ and $L_{\text{poms}} = 6 \times 10^3 L_\odot$ (2.3×10^{37} erg s⁻¹), one finds

$$T_e(R) = 3,530 R_{\text{AU}}^{-1/2} \left(\frac{L_{\text{poms}}}{6 \times 10^3 L_\odot} \right)^{1/4} \text{ K} \quad (2)$$

where R_{AU} is the comet's astrometric distance measured in AU. Equation (2) predicts that vigorous water sublimation will occur at distances up to ~ 400 AU around a typical post-main-sequence star. More limited water sublimation will proceed at distances as great as ~ 600 AU; CO and other extremely volatile ices will sublimate at distances as great as 10^4 AU.

The presence of a vast cloud of comets orbiting the Sun at large heliocentric distances is now well established^{2,3}. This assemblage of water-rich objects is believed to consist of three components. The innermost component of the Sun's comet cloud is the Kuiper (or Kuiper-Duncan) disk, a belt- or disk-like ensemble of $10^7\text{--}10^9$ cometary objects with an unknown amount of smaller debris. The Kuiper disk is believed⁴ to be the reservoir of the short-period comets, to lie between thirty and a few hundred AU from the Sun, and to have formed essentially *in situ* during the epoch of outer Solar System planetary formation. Various workers⁵⁻⁷ have demonstrated the presence of particle

disks at similar distances around other stars, but the presence of macroscopic objects (such as comets) in these disks has not been established. At a distance of $\sim 1,000\text{--}20,000$ AU from the Sun is the inner Oort cloud, a larger, more flared disk believed to consist of some $10^{12}\text{--}10^{13}$ comets ejected from the planetary region during the accumulation of the giant planets. The disk-like nature of the Kuiper Disk and inner Oort cloud result from their association with the planetary Solar System; dynamical (for instance, stellar) perturbations are insufficient to convert these inner disks into spheres on timescales comparable to the Sun's age. Finally, beyond the inner cloud is an isotropic outer Oort cloud containing $\sim 10^{11}\text{--}10^{12}$ comets extending as far as 2×10^5 AU from the Sun; the outer cloud is unstable to external perturbations and is resupplied from the inner cloud by the perturbing influence of passing stars and molecular clouds.

The thermal conditions described by equation (2) suggest that comets and other objects in the Kuiper disk will suffer intense mass loss during the Sun's post-main-sequence phase. In the more distant Oort cloud, comets on approximately circular orbits will be subject to very low water sublimation rates, but will experience some loss of volatile species including CO₂, CO, CN, Ar, N₂, NH₃ and CH₄ from their outer layers. As we will show, Oort cloud comets on eccentric orbits penetrating the sublimation region will suffer repeated, intense episodes of water sublimation and mass loss.

Sublimation of the comet cloud

When a solar-type star exhausts its hydrogen supply and ends hydrogen-core burning, its core contracts and hydrogen-shell burning begins. This causes the luminosity to steadily increase to a value of $\sim 300 L_\odot$ over $\sim 10^8$ years as the star becomes a red giant. At the end of this period, the water sublimation radius will have increased from 2.5 AU ($L=L_\odot$) to ~ 40 AU ($L=300 L_\odot$). This initial slow luminosity increase will have drastic

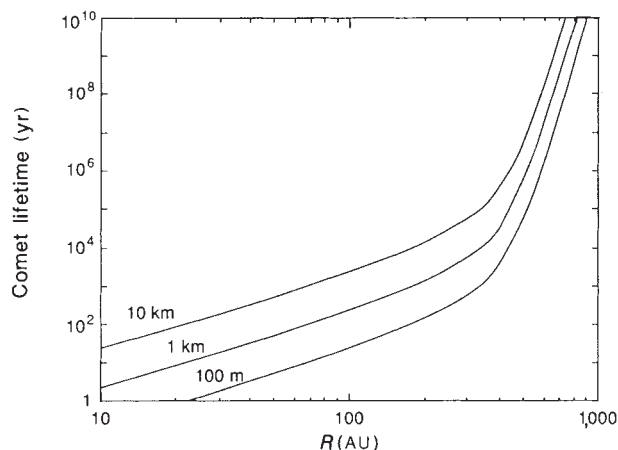


FIG. 1 Sublimation-destruction curves (equation (5)) for water-ice bodies of radius 100 m, 1 km and 10 km as a function of astrometric distance R for $L_{\text{poms}} = 6 \times 10^3 L_\odot$ and assuming a density of 0.5 g cm^{-3} .

and potentially observable influences on any planets and their satellites (a subject beyond the scope of this work). Because the volume of the sublimation zone increases in proportion to $L^{3/2}$, the number of active comets should increase dramatically.

Following complete hydrogen exhaustion and core contraction, the star will rise along the giant branch of the H-R (colour-luminosity) diagram, terminating in the helium flash. Although the flash luminosity may exceed $10^5 L_\odot$, the duration of the flash is too short (hours to days) to sublimate a large fraction of the cometary reservoir. The star then becomes a horizontal-branch star. For a period of $\sim 10^7$ years, low-mass post-main-sequence stars obtain a luminosity of several thousand $L_\odot$. During this period the water sublimation radius will extend into the region of the Kuiper disk and the inner Oort cloud.

To model the survival time of a comet disk against post-main-sequence heating, we conservatively assume that comets consist of only water ice and refractory dust. To calculate the water sublimation rate from a spherical icy body, we simultaneously solve for the global average surface temperature, including the effects of latent heat release

$$T(R) = \left[\frac{(1-A)L_{\text{poms}}/(16\pi R^2) - H(T)\dot{m}(T)}{\epsilon\sigma} \right]^{1/4} \quad (3a)$$

where the sublimation mass loss rate per unit surface area is

$$\dot{m}(T) = P_s(T) \sqrt{\frac{\mu}{2\pi kT}} \text{ g cm}^{-2} \text{ s}^{-1} \quad (3b)$$

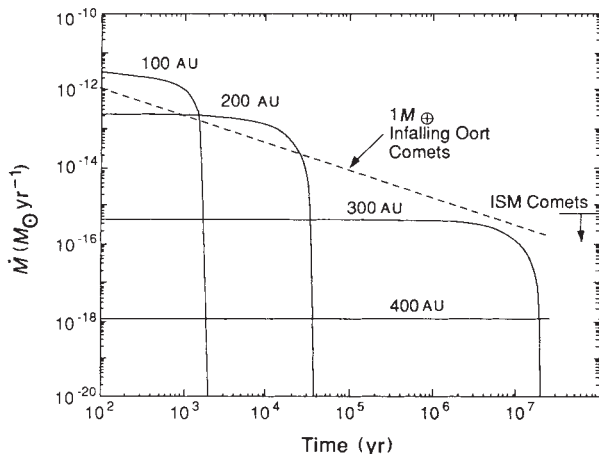
Here $H(T)$ is the specific heat of sublimation, $P_s(T)$ is the vapour pressure of water ice at temperature T , μ is the relative molecular mass of the water in grammes, and k is Boltzmann's constant. We use the relation for the saturation vapour pressure of water ice given by Lebofsky⁸

$$P_s(T) = 10^{-2.403.4/T+9.183837} \text{ torr} \quad (4)$$

where the temperature is specified in kelvin. Given a numerical solution to the mass loss rate, the lifetime of a spherical comet of radius r_c and density ρ at astrocentric distance R under constant stellar insolation is just

$$t_c(R, r_c) = \frac{4r_c\rho}{\dot{m}(R)} \quad (5)$$

Figure 1 depicts the sublimation timescale for water-ice objects, as a function of astrocentric distance. These curves were calculated according to the sublimation model given above by numerically shrinking test objects until they sublimated to destruction. Comets smaller than 10 km in radius and within 100 AU of such a star will be sublimated to destruction within $\sim 10^3$ years after a solar-type star becomes a luminous red supergiant. By the end of the post-main-sequence phase ($\sim 10^7$ yr), sublimation will have slowly destroyed the comets and other volatile-rich debris trapped inside a few hundred AU of the parent star.



In addition to the destruction of comets trapped on orbits entirely within the sublimation zone, infalling Oort cloud comets that pass through the sublimation zone will also lose mass. Models of the Oort cloud indicate that moderate orbital eccentricities are common, so this source of mass loss may be significant. For example, integrating $\dot{m}$ from equation (3b) over an orbit with a semi-major axis a of 1,200 AU and eccentricity e of 0.80, one finds that a 10-km comet consisting of 50% water ice by mass and having a 3% surface albedo will sublimate to destruction in 10^7 years around a $6 \times 10^3 L_\odot$ post-main-sequence star. On core helium exhaustion, solar-type stars form a planetary nebula and a white dwarf, which then fades slowly in luminosity.

We now explore the time-dependent evolution of comet disks and Oort-like clouds in the presence of a luminous post-main-sequence star. The solid lines in Fig. 2 depict the results of a sublimation rate calculation of mass loss in the Kuiper disk. An assemblage of comets of 5-km radius with total mass $1 M_\oplus$ were placed in a disk with a flat radial mass distribution extending from 100 to 1,000 AU. All objects were assumed to lie in nearly circular orbits in the disk. Each solid line in Fig. 2 represents the mass loss from those comets in a zone 100 AU wide centred on the labelled distance. The mass loss rates are for water ice; dust or other materials liberated by water sublimation will suffer loss rates in direct proportion to their relative abundance to water. In this model, each comet is diminished in size as it loses mass until it sublimates to destruction. Comet mantling effects were not included in the model. A constant $L_{\text{poms}} = 6 \times 10^3 L_\odot$ was assumed; higher luminosity results in a larger sublimation destruction zone, faster destruction, and increased mass loss rates at all distances. In a population of objects with various sizes, the smallest objects will be lost first; the instantaneous mass loss rate of the disk itself is not, however, set by the object sizes but by stellar luminosity and the total projected surface area of all objects in the disk.

Figure 2 shows that in the first $\sim 5 \times 10^4$ years after the post-main-sequence phase begins, sublimation will cause all comets trapped within 200 AU to sublimate to destruction. Beyond 300 AU, temperatures are sufficiently low that comets are not destroyed, even over 10^7 years. Thereafter, ignoring stellar pulsations, the loss rate of material from disk comets will consist of: (1) the partial destruction of those comets trapped within a few hundred AU of the star; (2) the sublimation of infalling comets that penetrate the water sublimation radius on orbits from the Oort-type cloud; and (3) the sublimation of any interstellar comets that pass through the sublimation zone of the star. The water mass loss rate, $\dot{M}_{\text{H}_2\text{O}}$, owing to the second mechanism (dashed line) was calculated for an orbit-averaged loss rate of a population of comets distributed uniformly in semi-major axis and eccentricity, as

$$\dot{M}_{\text{H}_2\text{O}} = \sum_a \sum_e N_c(a, e) \left(\frac{1}{T_{\text{orb}}} \right) \times \int_0^{T_{\text{orb}}} \pi r_c^2 \dot{m}_i(t, a, e, L_{\text{poms}}) dt \quad (6)$$

FIG. 2 Mass loss rate and survival time of a cometary disk around a $6,000 L_\odot$ star as a function of astrocentric distance. Each solid curve represents the mass loss rate from a 100-AU-wide annulus of the disk centred on the labelled distance. Dashed line gives the total (that is, distance integrated) mass loss rate from infalling comets derived from an Earth-mass ($1 M_\oplus$) Oort cloud. The horizontal bar on the right gives an upper limit for infalling interstellar comets.

Here $N_c(a, e)$ is the number of comets with orbital semi-major axis a and eccentricity e , and $\dot{m}_i(t, a, e, L_{\text{poms}})$ is the water-ice mass loss rate per unit surface area at time t in the orbit. T_{orb} is the orbital period for comets with semi-major axis a , circling a star of mass M and constant luminosity of $6,000 L_{\odot}$. For the case depicted in Fig. 2, comets of 5-km radius and total mass $1 M_{\odot}$ were distributed uniformly in 400 (a, e) orbit bins in the cloud.

For the case depicted (a $1 M_{\odot}$ of comets in the disk and equal mass in the outer Oort cloud, infalling comets dominate the mass loss after $\sim 5 \times 10^4$ years. The mass loss rate is proportional to the initial mass supply; therefore, if the mass of Oort cloud comets exceeds the mass of Kuiper disk comets, then the infalling Oort cloud comets can dominate the mass loss rate at earlier and earlier epochs. If the comet cloud has total mass of ~ 1 Jupiter-mass, then the dashed line would be raised by a factor of ~ 300 over what is shown.

An upper limit to the mass loss rate from interstellar comets passing through the sublimation zone on hyperbolic orbits was determined by calculating the sublimation rate for an interstellar comet population based on the upper limit⁹ on the space density of interstellar comets; the depicted result assumes a uniform distribution of impact parameters. We find that the expected mass loss from interstellar comets is small compared with that from the Kuiper disk and infalling comet cloud sources.

Comparing the total cometary mass loss over the post-main-sequence phase to the initial mass of the cometary reservoir, we found that although a substantial fraction (such as 30%) of the mass of a Kuiper-like disk can be entirely destroyed by post-main-sequence sublimation, only a small fraction (typically $\sim 1\%$) of the mass of the infalling Oort-like reservoir will be lost. Most objects in the halo are protected by their large periastron distances.

Because Oort clouds and comet disks are optically thin, the destruction timescale for any given object in the cloud is only a function of its albedo, mass and astrometric distance, and of the parent star's post-main-sequence luminosity. Smaller objects are destroyed faster. The time at which the mass loss curves in Fig. 2 turn down is set by the point at which the disk mass supply is exhausted at that distance. Therefore, the mass loss rate is supply-limited, rather than energy-limited. Further, although the disk mass loss rate at each distance depends on the total surface area exposed to insolation, the time at which the mass supply is exhausted is independent of the number of objects in the disk.

Cometary water in post-main-sequence outflows

The destruction of orbiting comets and debris provides a natural explanation for the water in, and several other attributes of, many post-main-sequence stellar outflows. These include the ubiquitous H_2O and OH emission surrounding post-main-sequence stars^{10,11}, the presence of HCN, organics and sulphur-bearing molecules around Miras and M supergiants¹² and the

coexistence of both ice grains and water at distances of several hundred AU in the outflows of red giants. Indeed, the water mass loss rates for many Mira and OH/IR star outflows^{11,13} are comparable to the range of sublimation rates predicted by our comet-cloud sublimation model.

The historical explanation for the production of water in stellar outflows is photochemistry^{14,15}. In these models, the OH abundance in circumstellar envelopes is controlled by exchange reactions and the dissociation of H_2O by interstellar ultraviolet radiation. The reaction $\text{OH} + \text{H}_2 \leftrightarrow \text{H}_2\text{O} + \text{H} + 0.69 \text{ eV}$ converts OH into H_2O in the warm inner envelope, whereas ultraviolet photodissociation produces OH in the outer envelope. These processes can explain OH densities $\sim 1 \text{ cm}^{-3}$ at distances of 10^{16} cm from the star¹⁴, which are comparable to those found in our models of Jupiter-mass Oort clouds during the first several hundred thousand years of post-main-sequence evolution (Fig. 3b). We believe that the comet-cloud sublimation model may provide a more natural, and in some respects more comprehensive, explanation for the OH and H_2O outflows.

We stress that the comet sublimation model does not predict the mechanism for H_2O or OH photon emission, only the ultimate source of the water. Presumably, the OH is created through dissociative kinetic and photon reactions of the cometary water vapour. Although post-main-sequence sublimation of the comet cloud can produce water mass loss rates comparable to those observed in the outflows of many Mira and OH/IR stars, we also point out that comet-cloud sublimation can account neither for the bulk stellar mass outflow rate nor for most of the dust in the outflow. That is, the sublimation mechanism is not a substitute for intrinsic stellar mass loss or for the production of dust in cool stellar atmospheres.

To demonstrate the inability of comet-cloud sublimation to provide the observed dust, we calculate the optical depth of a spherical dust outflow from the parent star owing to cometary sublimation. The dust:ice mass ratio in comets is $\sim 1:1$ so it is reasonable to assume that the dust and water production rates are similar. The dust optical depth as a function of wavelength λ can be written

$$\tau_{\lambda} \approx 10^{-3} \left(\frac{\dot{M}_c}{10^{-12} M_{\odot} \text{ yr}^{-1}} \right) \left(\frac{10 \text{ km s}^{-1}}{v_w} \right) \times \left(\frac{100 \text{ AU}}{R_s} \right) \left(\frac{0.2 \mu\text{m}}{\lambda} \right)^{1.5} \left(\frac{0.1 \mu\text{m}}{\bar{r}_g} \right) \quad (7)$$

where $\dot{M}_c$ is the dust mass loss rate assuming a dust:ice ratio of 1:1. Here v_w is the velocity of the stellar wind, R_s is the length scale of the sublimating region, and $\bar{r}_g$ is the number-weighted mean radius of the dust grains. Equation (7) assumes spherical cometary grains with $\rho = 0.2 \text{ g cm}^{-3}$ and an infrared emissivity of $(0.2 \mu\text{m}/\lambda)^{1.5}$. Using the $\dot{M}_c$ calculations shown in Fig. 2, equation (7) demonstrates that if an Earth-mass cometary disk sublimates, very low optical depths are to be expected. Compared with the nominal optical depths of OH/IR and Mira

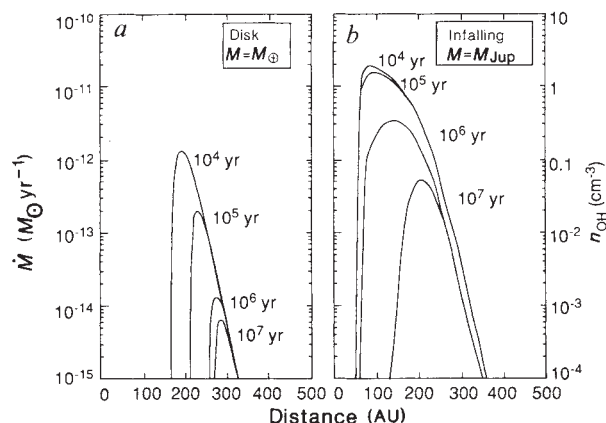


FIG. 3 *a* and *b* depict the mass loss rate of a Kuiper-like disk and a distant Oort-like cloud, respectively, around a $6,000 L_{\odot}$ star. Each solid curve represents the mass loss rate at the indicated time after the stellar luminosity is turned on. The strong dependence of sublimation rate on temperature produces an annular 'sublimation ring'. Variations in the stellar luminosity will cause the time evolution of the ring to be more complicated. The effect of infalling Oort cloud comets is to broaden the ring inward and prevent the peak mass loss rate from decaying below some level set by the orbital distribution and total mass of the infalling comets.

stars¹⁰, such optical depths are almost negligible.

Returning to water sublimation, an important distinguishing characteristic of a sublimating comet disk is its annular geometry. Such an annulus arises from the fact that water sublimation rates are extremely sensitive to temperature, and are therefore steep functions of the astrocentric distance. Figure 3a depicts the time evolution of such an annular ring (or torus) for a constant post-main-sequence luminosity of $6 \times 10^3 L_{\odot}$. Figure 3b depicts the mass-loss evolution from an infalling Oort-like halo of comets. The model for comet-disk sublimation produces geometric structures similar to actual very-long-baseline interferometry (VLBI) observations of some post-main-sequence stars. These observations^{13,16–18} show water and OH emission in smoothly expanding, thin 'shell-like' or 'ring-like' regions near the sublimation distance of 70–1,000 AU from the parent star (depending on the stellar luminosity). Although both the classical photochemical models and the comet-cloud sublimation model can explain shell-like haloes of water vapour, photochemical models cannot easily explain the annular geometries^{13,19}. Such geometries would naturally result from the sublimation of assemblages similar to the planar Kuiper disk.

The connection between cometary sublimation and the H₂O/OH outflows observed around M giants, Miras and OH/IR stars can be validated through future tests. For example, in the case of the ring-like emission regions, the sublimation model predicts that peak H₂O and OH number densities will lie inside the sublimation radius and near the rotational plane of the parent star. Furthermore, for both disk destruction and infall of Oort-type clouds, the model predicts an inverse correlation between the post-main-sequence age of the star and the water production rate, and a positive correlation between stellar luminosity and water mass loss rate. The model also predicts a phase-coherent correlation between the abundances of OH and H₂O and fluctuations in the bolometric luminosity of the parent star. One expects a phase delay between sublimation and stellar luminosity pulsations to be set by the time of light travel from the star to the location of the water source. In the complicated near region, such a phase delay will be short; at the far region, up to a distance of several hundred AU, the delay should be of the order of hours to days.

Finally, we suggest that the complicated, time-varying OH- and water-maser structures within 10–100 AU of many post-main-sequence stars, called 'blobs'^{13,16,19,20}, may actually represent individual orbiting bodies on eccentric trajectories. We note that H₂O and OH masers have been observed in cometary atmospheres²¹. According to stellar-atmosphere models, the post-main-sequence emission blobs are difficult to explain systematically and would not be expected to be in orbit around their parent stars. Although these emission blobs (or 'hotspots') have been observed to change position with time, no systematic study of their movements has been made. To test the comet-cloud hypothesis, we therefore suggest that these emission features should be observed (and tracked) on a timescale that is short compared with nominal keplerian timescales around Mira and OH/IR stars. By using both the high plane-of-sky spatial resolution and high line-of-sight velocity resolution inherent to VLBI, it should be possible to determine if the movements of these objects can indeed be fit to bound keplerian orbits, as would

be expected for comets. We point out that if the blobs actually move in keplerian orbits, then orbital inversion also offers the promise of accurately determining the masses of numerous single Miras and other post-main-sequence stars.

Summary and outlook

We have modelled the destruction of volatile-rich comet disks and Oort-type clouds around luminous post-main-sequence stars. Such models are in agreement with several aspects of existing observations of water and complex molecules in the envelopes of giant and supergiant stars. We suggest that these attributes result from the destruction of orbiting comet clouds. If confirmed, our results would establish the common existence of Oort-type clouds around other stars and would constitute indirect evidence for sites of past planetary formation. Struck-Marcell²² has conducted a study of the fate of planets and planetary satellites surrounding post-main-sequence stars. As shown there, high-luminosity phases of post-main-sequence stellar evolution can also have dramatic effects on planets.

We are now constructing more-detailed models of comet-cloud destruction and its observable signatures using numerical, time-dependent evolutionary tracks of post-main-sequence stellar evolution. In a future paper we will present these results and compare the applicability of the comet-destruction model to observations of specific well-known Miras and OH/IR stars.

For now, we have shown that luminous post-main-sequence stars will destroy resident volatile cometary-debris disks within several hundred to a thousand AU of the parent star on timescales shorter than the stellar post-main-sequence lifetime. Such a fate will remove the Solar System's Kuiper comet disk. Furthermore, comets on plunging orbits from the Oort cloud will repeatedly enter the post-main-sequence sublimation zone and lose substantial mass. The relative mass loss rates from the disk and infalling comet sources depend on the relative masses and orbital distributions of the disk and Oort cloud, and on the stellar luminosity. It is difficult to imagine a process that would quench comet-cloud sublimation around a luminous post-main-sequence star. Therefore, if volatile-rich bodies are orbiting such stars, high water production rates would naturally result.

The expected water loss rates from comet disks of several Earth masses or greater are similar to the observed mass loss rates around many post-main-sequence stars. Several other observable attributes of sublimating comet clouds are consistent with observations of many Mira and OH/IR stars. The Mira and OH/IR states are thought to be normal phases in the post-main-sequence evolution of late-type stars.

The importance of the detection of comets around stars is that comet formation strongly suggests planetary formation. Several observable signatures from sublimating comet clouds should be seen around red giants, Miras and other post-main-sequence stars possessing comet disks or clouds. We suggest that several already-observed attributes of Mira and OH/IR stars may be related to comet sublimation.

If comets are indeed responsible for commonly observed features in post-main-sequence stellar environments, then the probability of comet formation (and by extension, planetary systems) around single main-sequence stars is likely to be high. □

Received 2 January; accepted 9 April 1990.

1. Iben, I. & Renzini, R. *Phys. Rep.* **105**, 329–367 (1984).
2. Hills, J. G. *Astr. J.* **86**, 1730–1740 (1981).
3. Bailey, M. E., Clube, S. V. M. & Napier, W. M. *Vistas Astr.* **29**, 53–180 (1986).
4. Duncan, M., Quinn, T. & Tremaine, S. *Astrophys. J.* **328**, L69–L74 (1988).
5. Aumann, H. H. *et al.* *Astrophys. J.* **278**, L23–L28 (1984).
6. Harper, D. A., Loewenstein, R. F. & Davidson, J. A. *Astrophys. J.* **285**, 808–812 (1984).
7. Backman, D. & Gillett, F. C. in *Cool Stars, Stellar Systems and the Sun* (eds Linsky, J. L. & Stencel, R. E.) 340–350 (Springer, New York, 1987).
8. Lebofsky, L. A. *Icarus* **25**, 205–217 (1975).
9. Sekanina, Z. *Icarus* **27**, 123–133 (1976).
10. Jones, T. J., Hyland, A. R., Wood, P. R. & Gatley, I. *Astrophys. J.* **273**, 669–680 (1983).
11. Bowers, P. F. & Hagen, W. *Astrophys. J.* **285**, 637–647 (1984).
12. Nejad, L. A. M. & Millar, T. J. *Mon. Not. R. astr. Soc.* **230**, 79–86 (1989).

13. Bowers, P. F., Johnston, K. J. & de Vegt, C. *Astrophys. J.* **340**, 479–502 (1989).
14. Goldreich, P. & Scoville, N. *Astrophys. J.* **205**, 144–154 (1976).
15. Netzer, N. & Knapp, G. R. *Astrophys. J.* **323**, 734–748 (1987).
16. Reid, M. J., Muhleman, D. O., Moran, J. M., Johnston, K. J. & Schwartz, P. R. *Astrophys. J.* **214**, 60–77 (1977).
17. Jura, M. & Morris, M. *Astrophys. J.* **292**, 487–493 (1985).
18. Herman, J. & Habing, H. J. *Astr. Astrophys. Suppl.* **59**, 523–555 (1985).
19. Johnston, K. J., Spencer, J. H. & Bowers, P. F. *Astrophys. J.* **290**, 660–670 (1985).
20. Lane, A. P., Johnston, K. J., Bowers, P. F., Spencer, J. H. & Diamond, P. J. *Astrophys. J.* **323**, 756–765 (1987).
21. Crovisier, J. *Astr. J.* **90**, 670–674 (1985).
22. Struck-Marcell, C. *Astrophys. J.* **330**, 986–991 (1988).

ACKNOWLEDGEMENTS. We thank C. Hansen, I. Iben Jr, G. Knapp, M. Sölc, R. Stencel and P. Weissman for useful discussions. S.A.S. and J.M.S. acknowledge support from NASA.

Time-resolved X-ray crystallographic study of the conformational change in Ha-Ras p21 protein on GTP hydrolysis

Ilme Schlichting*, Steven C. Almo^{†‡}, Gert Rapp[§], Keith Wilson[§], Kyriakos Petratos[§], Arno Lentfer[§], Alfred Wittinghofer*, Wolfgang Kabsch*, Emil F. Pai*, Gregory A. Petsko^{†‡} & Roger S. Goody*^{||}

* Abteilung Biophysik, Max-Planck-Institut für Medizinische Forschung, 6900 Heidelberg, FRG

† Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

§ EMBL Outstation, DESY, 2000 Hamburg, FRG

Crystals of Ha-Ras p21 with caged GTP at the active site have been used to investigate the conformational changes of p21 on GTP hydrolysis. The structure of the short-lived p21·GTP complex was determined by Laue diffraction methods. After GTP hydrolysis, substantial structural changes occur in the parts of the molecule implicated in the interaction with GTPase-activating protein. The trigger for this process seems to be a change in coordination of the active-site Mg²⁺ ion as a result of loss of the γ -phosphate of GTP.

THE gene products (p21 proteins) of the *ras* oncogene family are thought to be involved in a signal transmission pathway similar to those involving G proteins^{1–3}. Analogous with the latter systems, and with other GTP-binding proteins such as the polypeptide elongation factors, p21 seems to be in an 'active' state when complexed with GTP, and in an 'inactive' state when complexed with GDP, the product of the p21 GTPase activity. It is this reaction, or more exactly its activation by GTPase-activating protein (GAP), that is modified in transforming mutants of Ha-Ras p21 (refs 4–8). Thus, an understanding of the GTP hydrolysis mechanism, the accompanying protein structural changes and the way in which these are modified by—and modify—the interaction with GAP are crucial for clarifying the molecular basis of oncogenic transformation by mutant p21 proteins.

We have recently shown that the Ha-*ras* oncogene product p21 can be crystallized with a photolabile GTP derivative (caged GTP; Fig. 1) at its active site, and that after photolytic removal of the protecting group, hydrolysis of GTP to GDP occurs in the crystal at the same rate as with p21 in solution⁹. We describe here the use of such crystals to provide information on the structural changes accompanying GTP hydrolysis in this system. Whereas the structures of static states of the system could be determined using classical (monochromatic) crystallographic methods, the short lifetime of the p21·GTP state made it necessary to use the polychromatic Laue diffraction method (for reviews, see refs 10, 11).

Experimental strategy

The structural data reported refer to the guanine-binding domain of p21 (amino acids 1–166, that is, without the C-terminal 23

residues), which is easily crystallized^{12,13}. The approach used is summarized in Fig. 1. The structures of stable or quasi-stable complexes in this scheme were determined by monochromatic X-ray diffraction methods. This was not possible for the relatively short-lived complex between cellular p21 (p21c) and GTP. This structure was determined by the Laue method using synchrotron radiation ($\lambda = 0.7$ – 2.0 Å) after cleavage of the 2-nitrophenylethyl group in crystals of p21c·caged GTP, and before significant hydrolysis of GTP to GDP had occurred. The structure of the protein was also determined (in the same crystal) by the Laue method, 14 min after initiating the GTPase reaction by removal of the protecting group, that is, during GTP hydrolysis (Fig. 2).

Description of the structures

(1) General. The overall structures of the complexes analysed here, p21·caged GTP, p21·GTP and p21·GDP, are the same in terms of the general organization of their secondary structural elements (α -helices, β -sheets) and also identical to that reported for the p21·GPPNP complex (see below)^{13,34}.

(2) p21·GTP. One of the main reasons for determining the p21·GTP structure was to compare it with the structure of p21 complexed to GPPNP, a non-hydrolysable GTP-analogue^{13,34}. The interactions around the site of modification of the analogue, the β - γ bridging atom, are important for judging the relevance of the latter structure and for justifying higher resolution studies.

A data set of p21 was obtained under Laue diffraction conditions during the first 5 min after photolysis of the crystalline p21·caged GTP complex (Fig. 2). We expected that the principal species at the active site would be GTP. This is confirmed by the electron density map (Fig. 3b) which shows the nucleotide bound to the active site in a manner analogous to GPPNP and with three clearly discernible phosphate peaks. No electron density was detected that could be attributed to the cage group. The base-protein interactions that have been identified in the p21·GPPNP structure are present, as are the interactions of the protein with the phosphate groups^{13,34}. Figure 4 demonstrates the similarity between the structures of the p21·GPPNP and p21·GTP complexes, as exemplified in the glycine-rich phosphate-binding loop. There is a hydrogen bond between the amide hydrogen of Gly 13 and the β - γ bridging oxygen of GTP or the β - γ bridging nitrogen of GPPNP, in the respective structures. Because the energy of the hydrogen bond to the nitrogen in GPPNP is expected to be less than that to the oxygen in GTP, this may explain the weaker binding of GPPNP to p21 (by a factor of ~ 10). This interaction may also explain why replacement of the bridging oxygen by groups which cannot form hydrogen bonds (CH₂, CF₂; ref. 12 and unpublished observations) leads to an even greater reduction of affinity than replacement by the NH group.

The Mg²⁺ ion is coordinated to the nucleotide by the γ - and β -phosphate groups (pro-R oxygen) and to the protein by the

‡ Present addresses: Department of Cell Biology and Anatomy, Johns Hopkins University, Baltimore, Maryland 21205, USA (S.C.A.), and Rosensiel Center, Brandeis University, Waltham, Massachusetts 02254-9110, USA (G.A.P.).

|| To whom correspondence should be addressed.

TABLE 1 Cell parameters and data statistics for p21 nucleotide complexes

Nucleotides	p21c		p21(G12V)		p21c		p21(G12V)		p21c		p21(G12V)	
	<i>a</i> = <i>b</i>	Cell axis <i>c</i> (Å)	<i>a</i> = <i>b</i>	<i>c</i>	Reflections (total)		Reflections (unique)		<i>R</i> _{sym} [†]		Resolution (Å)	
Caged GTP	41.0	164.8	41.0	177.7	7,290	8,143	2,703	3,551	12.8	8.2	3.0	2.8
GTP	41.1	164.8	40.5	161.3	**	11,868	**	5,411	**	4.7	**	2.6
GTP/GDP*	41.1	164.0	—	—	**	—	**	—	**	—	**	—
GDP	40.4	156.5	40.7	156.0	8,172	12,440	2,731	4,201	7.7	7.1	2.8	2.8

The large change in *c*-axis on GTP hydrolysis probably arises from the change in the structure of the effector region (Fig. 5). **, See Fig. 2 legend.

* Laue data set obtained 14 min after photolysis.

[†] $R_{\text{sym}} = \sum_i \sum_j |I_{h_i} - I_{h_j}| / \sum_i \sum_j I_{h_i}$, where I_{h_i} are the intensities of symmetry related reflections occurring uniquely.

functional groups of Ser 17 and Thr 35. In the 1.35-Å structure of p21·GPPNP, two molecules of water are also coordinated to the metal ion, one of them providing a bridge to the carboxyl group of Asp 57 (ref. 34). Only the latter has been identified in the p21·GTP structure. But as Mg²⁺ is normally hexacoordinated, it is likely that there is a water molecule in this position that has not been identified.

As in the p21·GPPNP complex, the ε-amino group of Lys 16 interacts with an oxygen of the γ-phosphate group (not the one that is a ligand to Mg²⁺), and also, more weakly, with the β-phosphate group. The latter interaction was not mentioned explicitly in ref. 13, but is present on re-inspection of the 2.5-Å structure of p21·GPPNP and is conspicuous in the 1.35-Å structure. So, the β- and γ-phosphate groups are doubly coordinated with Mg²⁺ and the NH₃⁺ group of Lys 16.

As in the p21·GPPNP complex, the residues in the loop L4 (amino acids 59–65) seem to be able to adopt several conformations. In particular, one position of the carbamoyl group of Gln 61 is close to a water molecule (Wat 175 in the p21·GPPNP structure³⁴) which itself is close to the γ-phosphate. This water molecule is believed to be involved in the GTPase reaction³⁴. In contrast to the p21·GPPNP structure, the density of the first residues of the following α-helix in the p21·GTP structure is not well defined.

Recent evidence suggests that there should be a structural difference between the p21·GTP complex described here and the p21·GPPNP complex described previously^{13,34}. An isomerization reaction of the p21·GTP complex has been identified and this seems to be rate-controlling in the overall GTPase reaction¹⁴. As this step also occurs with GPPNP, it follows that the p21·GPPNP complex should correspond to the structure of p21·GTP immediately before the hydrolysis step. But the p21·GTP complex examined here using the Laue method should correspond to the complex before the conformational change that leads to production of the hydrolytically competent state, as it is assumed to hydrolyse GTP too rapidly for it to accumu-

late. Similar experiments to those described here, but using caged GPPNP, should lead to a detailed characterization of this isomerization. The structural change must be subtle, as it has not so far been identified by comparison of the p21·GPPNP and p21·GTP structures. But it may occur in the region of amino acids 61–68 (loop L4 and the start of α-helix α2), which cannot be compared because of bad density in the p21·GTP structure.

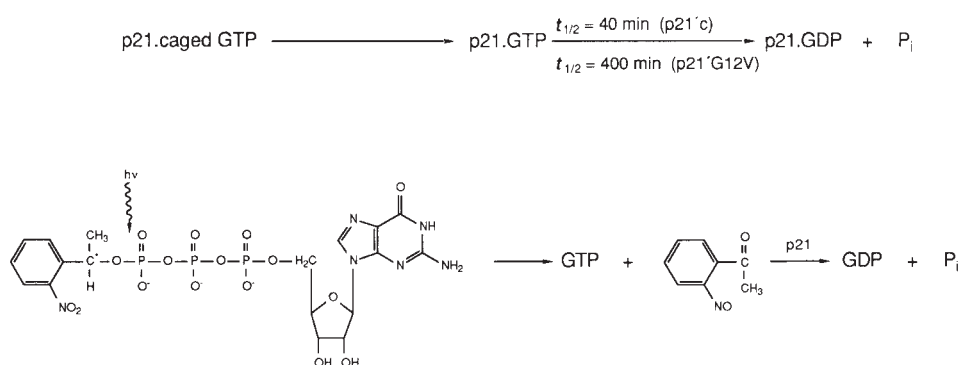
(3) p21·GDP. The structure of p21c·GDP (arising from hydrolysis of GTP after its photolytic release from caged GTP) was determined by crystallography using monochromatic radiation. There are several differences at the active site when compared with the GTP state, which are directly or indirectly due to the loss of the γ-phosphate group. Lys 16 still interacts with the backbone carbonyl group of residues 10 and 11, but not with the β-phosphate. The distance between the amino group of Lys 16 and the oxygens of the β-phosphate group is now too large for the interaction seen in the p21·GTP and p21·GPPNP states to occur. The coordination of Mg²⁺ is similar to that seen for p21·GTP or p21·GPPNP. But there is only one interaction with the nucleotide (β-phosphate group). The interaction with Asp 57 is now direct, as it is with the corresponding aspartate residue (Asp 80) in the structure of the EF-Tu·GDP complex^{15,16} and in the p21·GDP complex¹⁷, not through a water molecule as in the triphosphate structures. Also, the Mg²⁺ to Thr 35 distance has increased to a value that is too great for direct interaction (4.6 Å, compared with 2.4 Å in the p21·GTP state).

In contrast to the p21·GTP structure, there is good density for loop L4. This differs from the report of very weak electron density for residues 61–65 (ref. 17). The diameter of loop L4 has increased compared with the p21·GTP structure and the structure in this region is now well defined.

There are large differences between the structures of the effector-binding region (residues 32–40) in the p21·GTP and p21·GDP states (discussed below).

(4) The effector-binding region. The most dramatic differences between the p21·GDP structure and that of p21·GTP or

FIG. 1 Experimental scheme. Flow diagram for the generation of p21·GTP and p21·GDP complexes. p21c, Cellular (wild-type) protein; p21(G12V), oncogenic mutant in which glycine 12 is replaced by valine. Crystals of p21·caged GTP (ref. 9) were used to generate the GTP and GDP complexes of p21 by photolytic removal of the protecting 2-nitrophenylethyl group and subsequent hydrolysis. The structures of p21 complexed with caged GTP or GDP were determined by monochromatic X-ray diffraction methods, as was that of the GTP complex of p21(G12V) which was kept stable during the measurement by cooling to 4 °C. This was not possible for the relatively short-lived complex between p21c and GTP. This structure was determined by the Laue method (Fig. 2). Crystals of p21·GDP (nucleotide content verified by HPLC) were obtained after photolysing p21·caged GTP crystals and waiting for the GTPase reaction to take place. For practical reasons, X-ray



diffraction data were collected 19 h after photolysis. Monochromatic X-ray data were collected and processed as described¹³. The unit cell parameters and data statistics of the different p21·nucleotide complexes are summarized in Table 1.

p21·GPPNP are seen in this region (main differences between residues 32–36). This is the part of the molecule that is thought to interact with GAP (as shown by mutational analysis; ref. 18); it is therefore of interest to observe this structural change. Tyr 32 and Pro 34 have flipped to the other side of the loop, as has Asp 33 (albeit in the opposite direction) (Fig. 5). Tyr 32 now

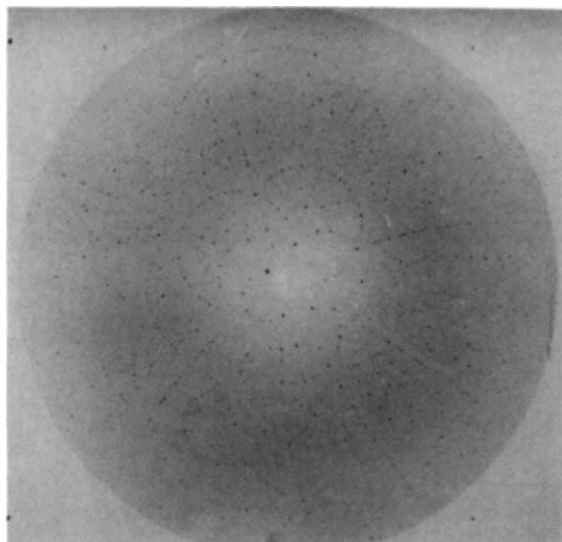


FIG. 2 Laue-Photograph of p21c·GTP obtained 4 min after photolysis. Crystals of p21·caged GTP (ref. 9; space group $P 3_2 2 1$, $a=b=41.0$ Å, $c=164.8$ Å, size $\sim 0.4 \times 0.4 \times 0.2$ mm) were mounted completely surrounded by artificial mother liquor by either wedging in tapered quartz capillaries or by immobilizing between pipe cleaner fibres. Only freshly prepared artificial mother liquor containing Tris buffer, $MgCl_2$, polyethylene glycol 400 and 40–50 mM DTE was used. Laue diffraction data were collected on film (one film pack consisting of six pieces of CEA reflex 25 film) at X31 located in Hamburg/DESY/HASYLAB during parasitic user time. For this Laue experiment, the white beam was focused by an eightfold segmented gold mirror with 1:1 focus, to obtain sufficient flux at the specimen. Absorption limited the long wavelength maximum to ~ 2.0 Å, whereas the short wavelength limit was 0.6–0.7 Å because the critical angle for reflection from the mirror was exceeded. The crystal was at random orientation relative to the X-ray beam, and the crystal–film distance was 11.8 cm. The angle of the spindle was tilted by 7° and exposures (10–15 s) of three different orientations of the crystal differing by 15° were collected for one data set. Only 2 out of 26 crystals of wild-type p21 were ordered sufficiently for Laue data analysis. Photolysis was achieved using a Xenon flash lamp⁹. Ten flashes, of ~ 1 min total, were used to obtain quantitative conversion to GTP (ref. 9). Laue data sets were collected from the same crystal 4 and 14 min after photolytic removal of the protecting group. A third data set 20 min after photolysis showed severe radiation damage (crystal was in the white beam for ~ 2 min) and thus was discarded. Films were scanned on an Optronics P-1000 HS rotation drum microdensitometer with a 50- μ m raster size and an optical density range of 0–2. Data were processed as described previously^{29,30} using a suite of programs (GENLAUE, INTLAUE, AFSCALE, LAUENORM) developed at Daresbury³¹. No wavelength deconvolution was used. There were no significant difficulties processing the Laue data collected in Hamburg with the programs optimized for the experimental conditions in Daresbury. During the refinement of the crystal orientation and the unit cell parameters (using the program GENLAUE; ref. 31) the crystal-to-film distance was fixed until the end of refinement. We discarded any questionable spots and therefore chose generous spot-to-spot distance constraints (1.0 mm) and, owing to the elongated diffraction spot shape (spot diameter was set to 0.8 mm in GENLAUE), a large box size (1.2 mm). The wavelength range used for scaling of the reflexes on the six films in one film-pack (AFSCALE; ref. 31) was 0.7–2.2 Å. This program also applies the obliquity and Lorentz-polarization correction. Only reflections with intensity $(I) > 1.5 \times \sigma(I)$ were retained for further processing. Finally, a wavelength-dependent normalization was performed with LAUENORM (ref. 31) by comparing Friedel equivalent reflections measured at different wavelengths. Wavelength curve fitting was performed with two separate curves to eliminate discontinuity associated with the absorption edge of bromine (0.97 Å) in the photographic films. As the additional gold absorption edge (due to coating of the mirrors) in the intensity versus wavelength-dependence coincides with the bromine absorption edge, it was not necessary to use more complicated wavelength normalization procedures. The best R -factor was obtained using a wavelength range 0.69–1.95 Å, giving 2,193 unique reflections from the total of 3,817 obtained from the AFSCALE output. The results of data reduction are summarized below.

	p21c·GTP	p21c·GTP/GDP
$R_{\text{merge}}(1, 2, 3)$ (see ref. 1)	7.8/7.6/7.2	7.4/7.2/5.6
Reflections (total)	3,817	4,338
Reflections (unique)	2,193	2,322
d (% of possible)	2.8 Å (50%)	2.8 Å (50%)

seems to interact more strongly with the pyrrolidone ring of Pro 34 than it does in the GTP structure. We cannot yet provide a detailed causal explanation for this structural change but it seems likely that the conformational change is associated with, and is perhaps due to, a change in the position of the metal ion. The interaction of Mg^{2+} with the (now missing) γ -phosphate group seems to be replaced in the p21·GDP structure by a direct interaction with the carboxyl group of Asp 57 (in p21·GTP and p21·GPPNP, this interaction is indirect through a water molecule). Thus, the double positive charge on the Mg^{2+} ion, that is neutralized in the triphosphate structures by interaction with two phosphate groups, is now neutralized by one phosphate group and one carboxyl group. In addition, as mentioned above, Thr 35 is no longer directly coordinated to Mg^{2+} in the p21·GDP state. The same structural changes are observed in the transforming mutant protein p21(G12V) (Gly 12 $\rightarrow$ Val). As this is known to interact with GAP in the triphosphate state, and not in the GDP-bound state, it is not surprising that the same structural change is observed on hydrolysis of GTP.

Milburn *et al.* have also compared the structures of the p21·GDP and p21·GPPCP complexes¹⁹. In agreement with our results, the main differences seen by this group are in the effector loop and in loop L4. Although some of these changes seem to be similar to those found in the present study, there are many differences in detail. This applies to both of the loops, and to the magnesium coordination. A thorough comparison of the results obtained will depend on the availability of coordinates for all of the relevant structures involved.

(5) p21·caged GTP—'incorrect' binding in the crystal due to steric hindrance and crystal packing forces. The binding of caged GTP to p21 in the crystal differs from that of GTP, GPPNP or GDP (Fig. 3a). None of the main interactions seen with the other nucleotides is seen with caged GTP. First, the base is too far from Asp 119 for the strong interactions with N1 and N2 present in the other structures to occur. Phe 28 is also incorrectly orientated for an interaction with the aromatic ring system. These two interactions are thought to be of prime importance for positioning (Phe 28, together with the aliphatic side chain of Lys 116) and discrimination (Asp 119) of the base from the other nucleotides. In the triphosphate part of the molecule, the ribose is at roughly the same position as in the other complexes, but the α -, β - and γ -phosphates are far removed. Because of this, the important interactions of the latter two groups with the protein and with magnesium cannot occur. There are interactions of the 3'-hydroxyl group with the carbonyl group of Val 29, and of the oxygens of the γ -phosphate group with the amide hydrogens of Tyr 32 and Asp 33. None of these interactions is seen in the complexes between p21 and GPPNP, GTP or GDP. There seems to be an aromatic interaction of the cage group with Tyr 32 and of the nitro group with the side-chain carbamoyl group of Asn 86 of the neighbouring p21 molecule. It is not known whether magnesium is bound in this complex, but it seems likely that electron density associated with Ser 17 arises from the magnesium ion. There are some differences in the binding of caged GTP to p21(G12V) in the crystal. The nucleotide is twisted even further from its position in the other complexes. In contrast to the p21c·caged GTP complex, there is an interaction between the amino group of Lys 16 and the β -phosphate group.

The unusual mode of binding of caged GTP in the crystals prompted a more detailed investigation of the kinetics of interaction of caged GTP with p21 in solution. We have previously reported that it binds about 100-fold more weakly than GDP (ref. 9). A more detailed study, using a stopped-flow kinetic method²⁰, has led to values for the association rate constant ($2.5 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$) and the dissociation rate constant ($2.5 \times 10^{-4} \text{ s}^{-1}$; both values at 25°C). The corresponding values for GDP are $8.5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ and $1.2 \times 10^{-5} \text{ s}^{-1}$, respectively²⁰. Thus, from the quotient of the rate constants, caged GTP binds only a factor of ten more weakly than GDP (that is, 10^{10} M^{-1}

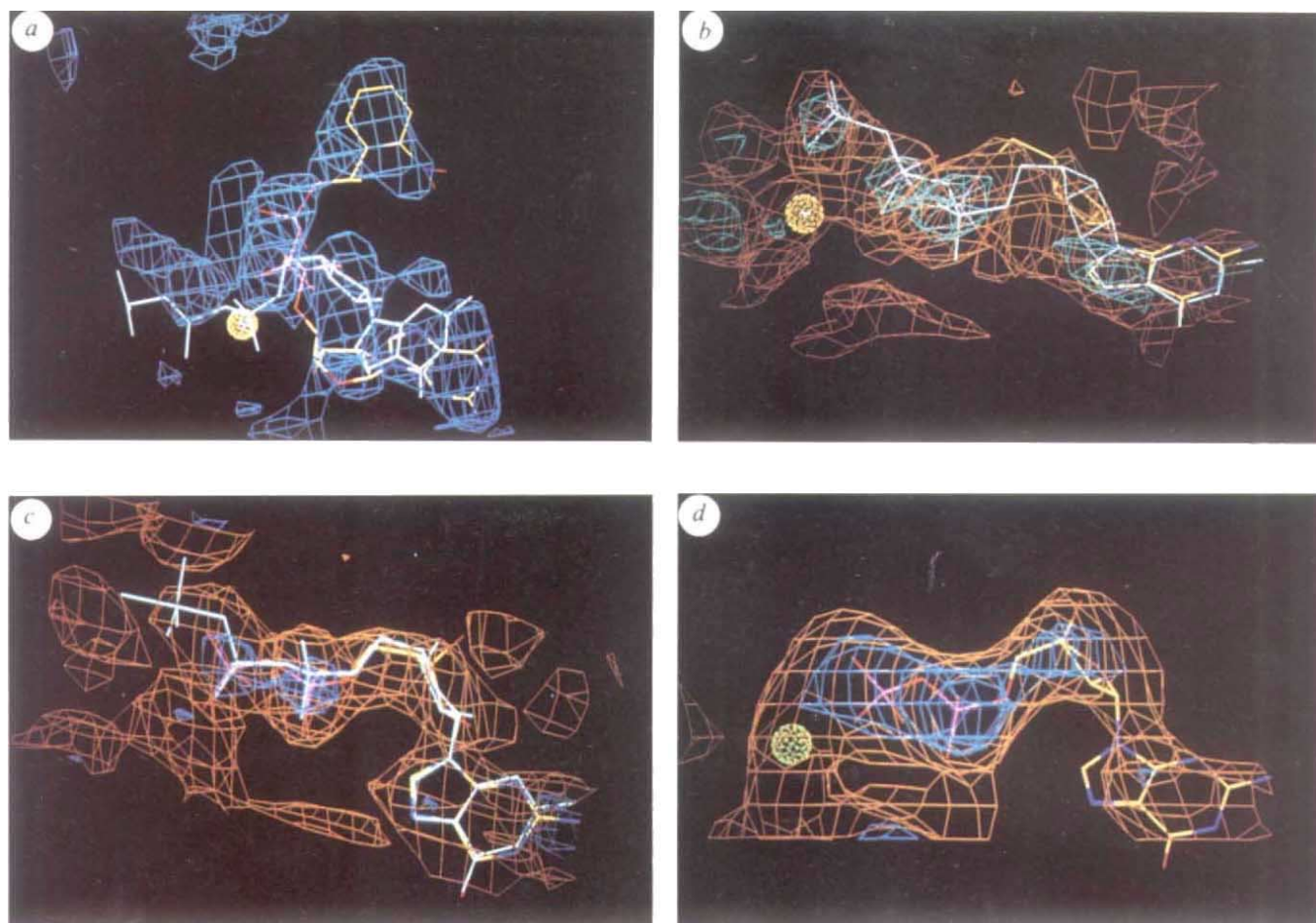


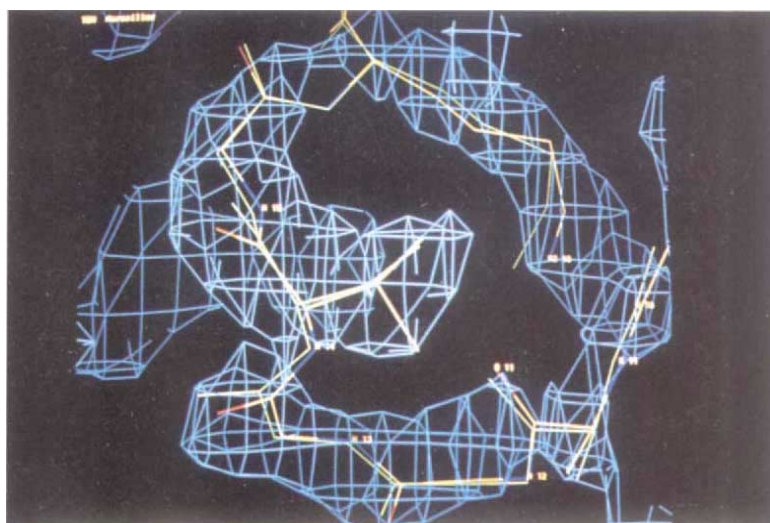
FIG. 3 Electron density of the bound nucleotide at the active site of p21c. *a*, Caged GTP. *b*, GTP, Laue data obtained 4 min after photolysis. *c*, GTP/GDP, Laue data obtained 14 min after photolysis. *d*, GDP. Electron density levels (blue, yellow respectively) are given as percentage of the maximum electron density: *a*, 45; *b*, 50, 22; *c*, 45, 10; *d*, 50, 22. Magnesium is shown as a sphere. In *a*, the structure of GTP (white) from *b* is superimposed on the caged GTP structure to show the different mode of binding of the latter. In *b*, the structure of GPPNP (ref. 13) is shown in white superimposed on the GTP structure. In *c*, the GTP structure (white) from *b* is superimposed on the GDP structure. It can be seen that there is much less density at the position of the γ -phosphate. The loss of density is greater than the $\sim 30\%$ expected from the rate constant for GTP cleavage. This effect could be caused by heating of the crystal in the white X-ray beam, which occurs at a significant rate. No other density that could be assigned to that missing from the terminal phosphate group was found. Refinement was performed

with XPLOR (ref. 32) using the coordinates of the p21·GPPNP structure as a starting model^{1,3}. About 100 water molecules were fitted in the p21c·GTP structure. Refinement parameters are given below.

	p21c				p21(G12V)		
	Caged GTP	GTP	GTP/GDP	GDP	Caged GTP	GTP	GDP
R_{ini}	46.2	30	40	46.8	53.7	33.6	41.8
N_{cyc}	4	6	1	3	1	2	2
R_{fin}	19.9	14.0	14	18.5	25.3	21.4	17.9

N_{cyc} represents the number of XPLOR runs, each consisting of simulated annealing and energy/X-ray refinement with manual rebuilding using FRODO (ref. 33) between runs. $2F_o - F_c$ electron density maps were calculated. R_{fin} is given in per cent (water molecules not included).

FIG. 4 Electron density (at 16% of the maximum value) in the p21c·GTP structure at the glycine-rich phosphate-binding loop. The good agreement between the structures of p21 complexed to GTP and GPPNP (in light green; ref. 13) is apparent.



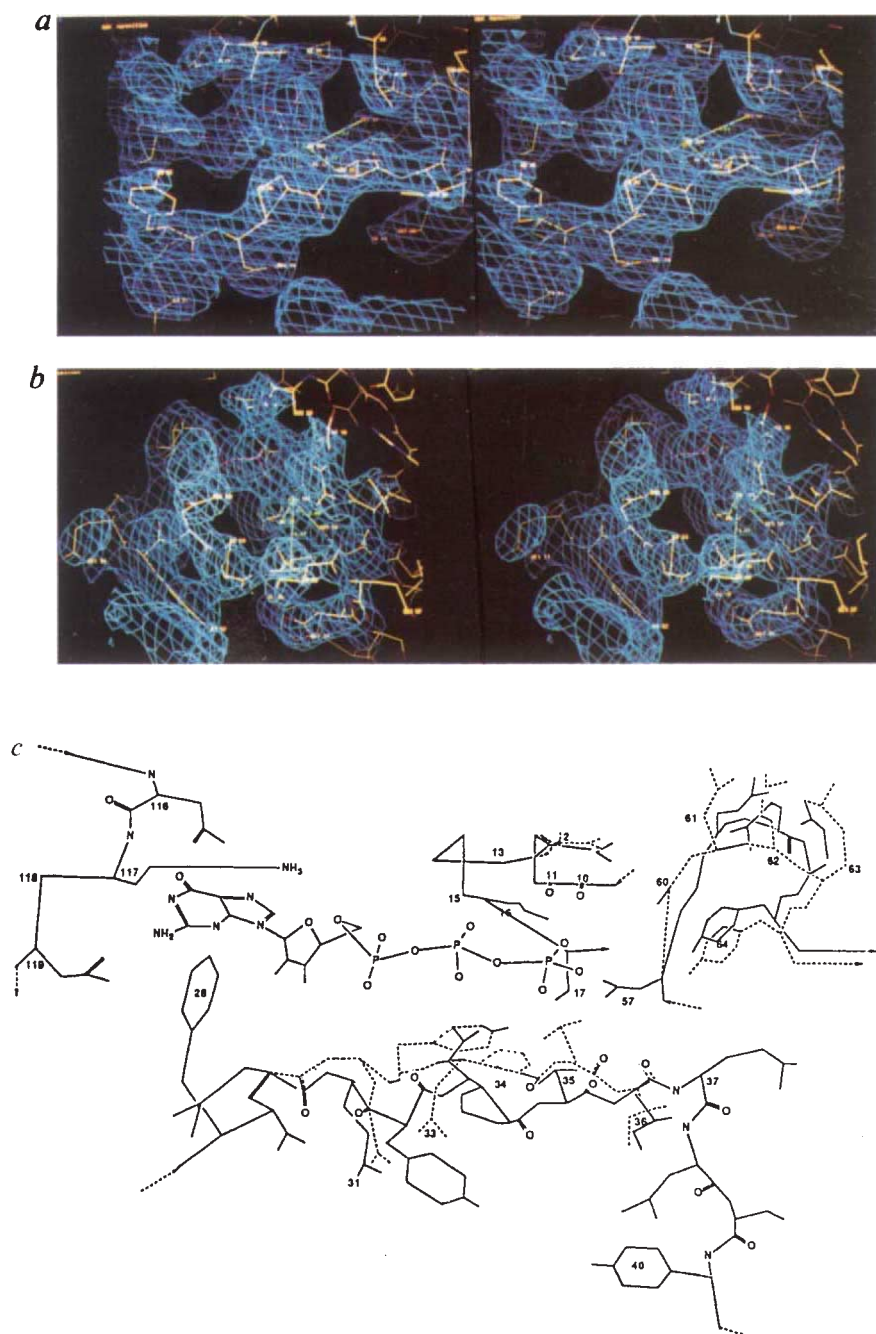
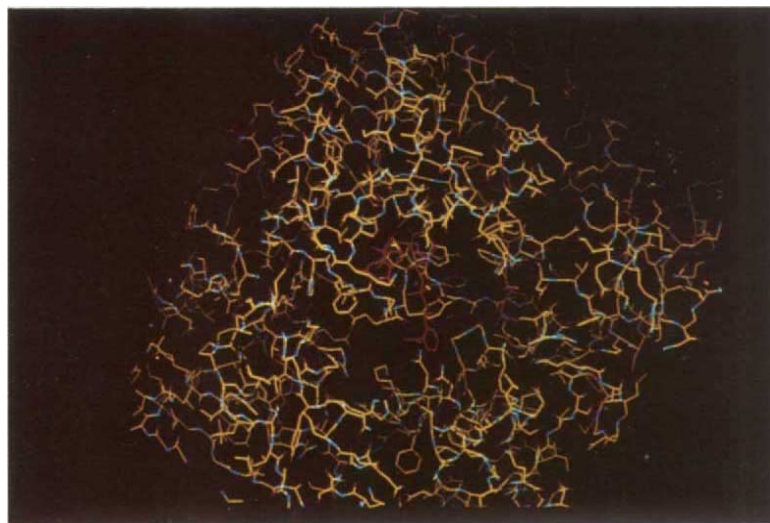


FIG. 5 Stereo view of the electron density (shown at 16% of the maximum value) of the effector region of p21(G12V) complexed to GTP (a) and GDP (b). There is a large movement of the residue 32–36 region between the two complexes. This is clarified in the schematic diagram c. The structure shown is the p21 · GTP complex, with the main changes in the p21 · GDP complex shown by the dashed lines.

FIG. 6 Binding of caged GTP. The protecting 2-nitrophenylethyl group points towards a solvent channel ($\sim 15 \times 20$ Å) parallel to the *a*-axis of the crystal lattice.



compared with 10^{11} M^{-1}). This is difficult to reconcile with the observed mode of binding of caged GTP in the crystal. In addition, as with GDP and GTP, there is a marked effect of Mg^{2+} on the dissociation kinetics of caged GTP from p21, which does not accord with the apparent lack of interaction of Mg^{2+} with caged GTP at the active site of p21. Further evidence for the 'correct' binding of caged GTP in solution comes from two-dimensional NMR measurements on the p21·caged GTP complex (data not shown). In particular, the distance between H1' of the ribose ring and Phe 28 is found to be $\sim 3.5 \text{ \AA}$ in solution (similar to GDP and GPPNP; ref. 21 and unpublished results), compared with $6 \text{ \AA}$ in the crystal structure. It is apparent from the structures seen after removal of the protecting group, that this altered mode of binding in the crystal does not cause any obvious problems for these experiments. In future experiments the unusual binding mode may even be advantageous. It could allow characterization of substrate-induced structural changes on correct binding of photogenerated GTP. We expect these to occur at a rate constant of $\sim 30 \text{ s}^{-1}$ at 25°C ($\sim 5 \text{ s}^{-1}$ at 4°C) from kinetic experiments in solution²⁰.

Modelling of caged GTP in the active site of p21 (keeping the GTP part of the analogue in the same position as in the p21·GPPNP or p21·GTP complexes) provides an explanation for these observations. It is seen that this would cause the protecting group to collide with the functional groups of Gln 61 and Tyr 32 of the neighbouring p21 molecule in the crystal structure. This seems to force the caged GTP molecule out of its preferred binding position, and indicates that the crystal packing energy must be greater than the binding energy of the analogue in its proper GTP-like position. In the new binding position, there is no steric hindrance of the protecting group (Fig. 6).

(6) Comparison of p21c and p21(G12V). The structures of p21c and p21(G12V) are very similar over most of the molecule in both the GTP and GDP forms. In particular, a similar conformation of the effector loop is seen in the respective GDP forms, and the same structural change occurs on hydrolysis to GTP. This similarity accords with the observations that the GTP form of p21(G12V) can bind to GAP with an affinity of the same order of magnitude as that of p21c, and that its affinity for GAP in the GDP form is much lower, as for p21c (refs 7, 8). There are, however, differences in the structures in the region of the phosphate-binding loop (although we do not observe the large widening of the loop as reported in ref. 22) and in the loop L4, the latter being pushed further away from the γ -phosphate group. This is particularly noticeable in the larger distance between the phosphate oxygens, the amide proton of Gly 60 and the carbonyl oxygen of Thr 58 on the one hand, and the side-chain of Gln-61 on the other hand. It is likely that these changes are associated with the reduced rate of the GTPase, particularly in the presence of GAP, in this mutant compared with p21c. The detailed analysis of transforming mutant protein structures at higher resolution than that available in the present study provides more information on these effects (U. Krengel *et al.*, manuscript submitted).

Discussion

Although it is not possible to draw detailed conclusions about the mechanism of GTP hydrolysis at the active site of p21, general features can be deduced from the available evidence. First, as is apparent both from this work and from the results on the structure of the p21·GPPNP complex, there are many interactions between the β - and γ -phosphate groups, the protein and Mg^{2+} . These interactions both increase the electrophilicity of the γ -phosphate group (and therefore its susceptibility to attack by a water molecule) and improve the leaving group properties of the β -phosphate. These features are presumably shared by many other enzymes that catalyse phosphate transfer, as many of these have the conserved phosphate-binding loop (residues 10–18) characteristic of such enzymes^{23,24}. Lys 16

seems to play a direct part in the catalytic mechanism. It is an invariant feature of the phosphate-binding loop and is always followed directly, or after an additional glycine, by a serine or threonine residue, in which p21 and presumably also in the other phosphate-transferring proteins, is complexed to Mg^{2+} . Preliminary modelling studies indicate that the interaction between the side-chain amino group of Lys 16 and the γ -phosphate group becomes even stronger in the putative pentacoordinate intermediate, formed after the attack of water at phosphorous known to be direct from stereochemical studies²⁵. The side chain of Lys 16 seems to be held in position by strong interactions with the main-chain carbonyl groups of residues 10 and 11. The structure envisaged would be very similar to the putative structure of transition-state adenylate kinase³⁵, ribonuclease A (ref. 26) and the complex between creatine kinase, ADP, a planar nitrate ion and creatine²⁷.

A water molecule has been identified in the p21·GTP structure presented here and the high resolution p21·GPPNP structure³⁴ that is likely to be the one that attacks the phosphate group. The similarity between these two structures and the explanation of the difference in the affinities of p21 for GTP and GPPNP justifies the derivation of biological and structural conclusions from the 121·GPPNP structure.

Some indications of the mechanism of activation of the attacking water molecule can be derived from the available structures (as mentioned already) and are discussed in more detail elsewhere³⁴. Further time-resolved studies using caged GPPNP, caged GTP (γ -S) and caged GTP should provide more information about this process.

Further studies are needed to clarify the structural change seen in the effector loop on GTP hydrolysis. Several technical improvements will help. First, changes in the nature of the photolytic protecting groups used could lead to structures at higher resolution for the various states characterized in this work. An obvious improvement is to use single diastereomers of caged GTP, which could lead to improvement of crystal quality, allowing time-resolved crystallography at higher spatial resolution. The use of a more powerful light source (laser) will improve the temporal resolution immediately after initiation of GTP hydrolysis. This could be complemented by use of a more sensitive detector system (probably imaging plates), that should lead to a reduction of at least a factor of ten in exposure time and increase (by a factor of ten or more) the number of data sets available from a single crystal. In this study, the highest number of data sets obtained before radiation damage became unacceptable was two. Another advantage is that the crystal heating by the white X-ray beam would also be reduced by a similar factor. A combination of these features, with increased X-ray intensity (available at several synchrotron sources) will extend investigations such as these into the sub-second and even sub-microsecond time range²⁸. The work described here demonstrates that photolytic generation of substrates at the active site of a protein in the crystalline state is a viable approach to rapid initiation of an enzymatic reaction, which is needed to take advantage of the potential of synchrotron radiation for time-resolved studies.

Coordinates of the p21·GTP structure from the 1.35- $\text{\AA}$ electron density map will be deposited in the Brookhaven Data Bank. □

Received 7 March; accepted 12 April 1990.

1. Barbacid, M. A. *Rev. Biochem.* **56**, 779–827 (1987).
2. Gilman, A. G. A. *Rev. Biochem.* **56**, 615–649 (1987).
3. Allende, J. E. *FASEB J.* **2**, 2356–2367 (1988).
4. McCormick, F. *Cell* **56**, 5–8 (1989).
5. Adari, H., Lowy, D. R., Willumsen, B. M., Der, C. J. & McCormick, F. *Science* **240**, 518–521 (1988).
6. Cales, C., Hancock, J. F., Marshall, C. J. & Hall, A. *Nature* **332**, 548–551 (1988).
7. Marshall, M. S. *et al.* *EMBO J.* **8**, 1105–1110 (1989).
8. Vogel, U. S. *et al.* *Nature* **335**, 90–93 (1988).
9. Schlichting, I. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **86**, 7687–7690 (1989).
10. Moffat, K. A. *Rev. Biochem.* **18**, 309–332 (1989).
11. Hajdu, J. & Johnson, L. N. *Biochemistry* **29**, 1669–1678 (1990).
12. Scherer, A. *et al.* *J. molec. Biol.* **206**, 257–259 (1989).

13. Pai, E. F. *et al. Nature* **341**, 209-214 (1989).
14. Neal, S. E., Eccleston, J. F. & Webb, M. R. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
15. Jurnak, F. *Science* **230**, 32-36 (1985).
16. Woolley, P. & Clark, B. F. C. *BioTechnology* **7**, 913-920 (1989).
17. De Vos, A. M. *et al. Science* **239**, 888-893 (1989).
18. Gibbs, J. B., Schaber, M. D., Allard, W. J., Sigal, I. S. & Scolnick, E. M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5026-5030 (1988).
19. Milburn, M. V. *et al. Science* **247**, 939-945 (1990).
20. John, J., Sohmen, R., Feuerstein, J., Wittinghofer, A. & Goody, R. S. *Biochemistry* (in the press).
21. Schlichting, I. *et al. Biochemistry* **29**, 504-511 (1990).
22. Tong, L., Milburn, M., DeVos, A. M. & Kim, S. H. *Nature* **337**, 90-93 (1989).
23. Gay, N. J. & Walker, J. E. *Nature* **301**, 262-264 (1983).
24. Leberman, R. & Egner, U. *EMBO J.* **3**, 339-341 (1984).
25. Feuerstein, J., Goody, R. S. & Webb, M. R. *J. biol. Chem.* **264**, 6188-6190 (1989).
26. Deakyn, B. & Allen, J. *Am. chem. Soc.* **101**, 3951-3959 (1979).
27. Kenyon, G. L. & Reed, G. H. *Adv. Enzym.* **54**, 367-426 (1983).
28. Szebenyi, D. M. E. *et al. Trans. Am. crystallogr. Ass.* **24**, 167-172 (1988).
29. Farber, G. K., Machin, P., Almo, S. C., Petsko, G. A. & Hajdu, J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 112-115 (1988).
30. Hajdu, J. *et al. Nature* **329**, 178-181 (1987).
31. Helliwell, J. R. *et al. J. appl. Crystallogr.* **22**, 483-497 (1989).
32. Brünger, A. T., Kuriyan, K. & Karplus, M. *Science* **235**, 458-460 (1987).
33. Jones, T. A. *J. appl. Crystallogr.* **11**, 268-272 (1978).
34. Pai, E. F. *et al. EMBO J.* (in the press).
35. Reinstein, J., Schlichting, I. & Wittinghofer, A. *Biochemistry* (in the press).

ACKNOWLEDGEMENTS. We thank Marija Isakov and Anna Scherer for technical assistance, Ken Holmes for support, encouragement and suggestions during the writing of the manuscript, Lynne Howell for help with data reduction, and Greg Farber and Barry Stoddard for help with the Laue experiments. G.A.P. thanks the Alexander von Humboldt Foundation for support. The work was also supported by the European Community.

Deficiency of a glycoprotein component of the dystrophin complex in dystrophic muscle

James M. Ervasti, Kay Ohlendieck, Steven D. Kahl, Mitchell G. Gaver & Kevin P. Campbell*

Howard Hughes Medical Institute and Department of Physiology and Biophysics, University of Iowa College of Medicine, Iowa City, Iowa 52242, USA

Dystrophin, the protein encoded by the Duchenne muscular dystrophy (DMD) gene, exists in a large oligomeric complex. We show here that four glycoproteins are integral components of the dystrophin complex and that the concentration of one of these is greatly reduced in DMD patients. Thus, the absence of dystrophin may lead to the loss of a dystrophin-associated glycoprotein, and the reduction in this glycoprotein may be one of the first stages of the molecular pathogenesis of muscular dystrophy.

DUCHENNE muscular dystrophy is caused by a defective gene located on the X chromosome. Dystrophin, the high-molecular weight protein product of the DMD gene¹, is localized to the sarcolemmal membrane of normal skeletal muscle²⁻⁵ but is absent from the skeletal muscle of people with DMD^{1,2,6}, *xmd* dogs⁷ and *mdx* mice^{1,5} (the last two being possible animal models for DMD). The amino-acid sequence of dystrophin suggests that it is a membrane cytoskeletal protein^{8,9} involved in the anchoring of sarcolemmal proteins to the underlying cytoskeleton. But the exact function of dystrophin and its precise role in the resulting necrosis of dystrophic muscle fibres has not been determined. In studies of other genetic diseases involving proteins of the cytoskeleton^{10,11}, the absence of one component is sometimes accompanied by the loss of another cytoskeletal protein. Therefore, to understand the molecular pathogenesis of DMD, we sought to identify the proteins associated with or bound to dystrophin and to characterize the status of these proteins in muscle where dystrophin is absent.

Recently, we have shown that dystrophin can be isolated from detergent-solubilized skeletal muscle membranes using wheat-germ agglutinin (WGA)-Sepharose, because of its tight association with a WGA-binding glycoprotein¹². This indicates that the localization of dystrophin to the cytoplasmic face of the sarcolemma²⁻⁵ results from a tight association of dystrophin with an integral membrane glycoprotein. Here we report the purification of a large oligomeric complex (~18S) containing

dystrophin using sucrose density-gradient centrifugation in the presence of digitonin. We have identified four glycoproteins of apparent relative molecular masses (M_r) 156,000 (156K), 50K, 43K and 35K as integral components of the dystrophin complex. The 156K and 50K glycoproteins are sarcolemmal glycoproteins, as shown by indirect immunofluorescence. Immunoaffinity beads raised against dystrophin and the 50K glycoprotein selectively adsorb the dystrophin-glycoprotein complex. Furthermore, there is a marked reduction of the 156K glycoprotein in muscle from *mdx* mice and DMD patients. These results imply that in dystrophic muscle, the absence of dystrophin may lead to the loss of a dystrophin-associated glycoprotein. This could be the first step in the molecular pathogenesis of muscular dystrophy.

Dystrophin-glycoprotein complex

This complex was isolated following digitonin-solubilization of rabbit skeletal muscle membranes using WGA-Sepharose and DEAE-cellulose¹² and further purified by sucrose density gradient centrifugation in the presence of 0.1% digitonin. It is evident from the Coomassie blue-stained gel of sequential gradient fractions (Fig. 1a) that the dystrophin-glycoprotein complex was separated from the voltage-sensitive sodium channel and the dihydropyridine receptor (Fig. 1). The size of the dystrophin complex was ~18S in comparison with β -galactosidase (15.9S), thyroglobulin (19.2S) and dihydropyridine receptor (20S) standards. Densitometer scanning of the peak dystrophin-containing fractions (10 and 11, Fig. 1a) revealed several proteins that co-purified with dystrophin: a broad, diffusely staining component with an apparent M_r of 156K, an 88K protein, a triplet of proteins centred at 59K, a 50K protein, a doublet at 43K and proteins of 35K and 25K.

To identify the glycoprotein constituents of the dystrophin-glycoprotein complex, sucrose gradient fractions 7-17 were electrophoretically separated, transferred to nitrocellulose and stained with peroxidase-conjugated WGA (Fig. 1b). Four WGA-binding proteins with apparent M_r of 156K, 50K, 43K and 35K were found to strictly co-purify with dystrophin. All four proteins were also positively stained with peroxidase-conjugated concanavalin A. In addition, the lower M_r component of the 43K protein doublet (Fig. 1a) was also positively stained with concanavalin A (not shown).

The dystrophin-glycoprotein complex was further characterized with antibodies raised against various components of the

* To whom correspondence should be addressed.

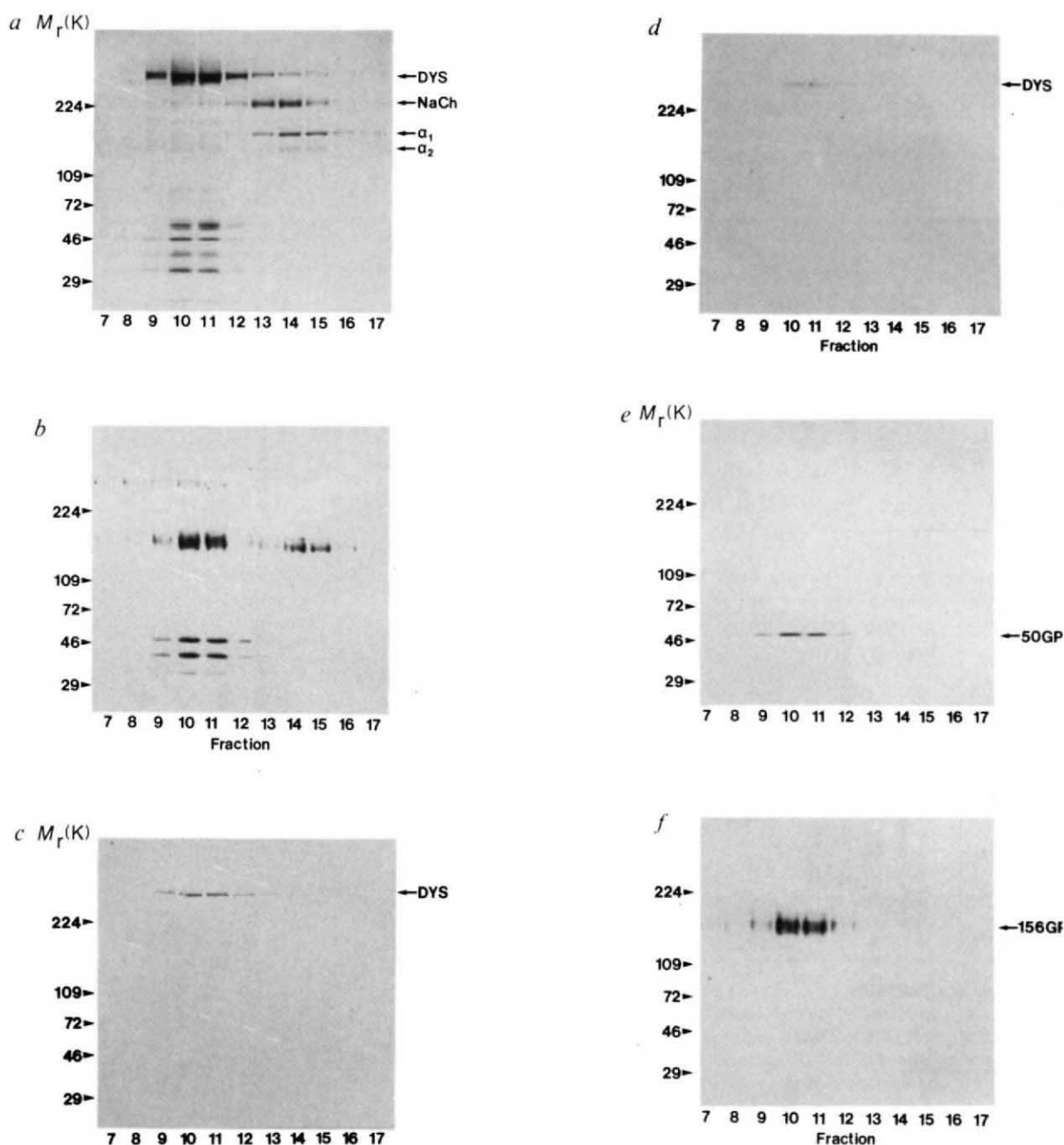


FIG. 1 Sedimentation of dystrophin complex through 5% to 20% linear sucrose gradients. *a*, Coomassie blue-stained gel of sucrose gradient fractions 7–17. *b–f*, Nitrocellulose transfers of sucrose gradient fractions 7–17 separated by SDS-PAGE and stained with: *b*, peroxidase-conjugated WGA ($1 \mu\text{g ml}^{-1}$); *c*, polyclonal antisera against the C-terminal decapeptide of dystrophin; *d*, monoclonal antibody (mAb) XIXC2 against dystrophin; *e*, mAb IVD31 against 50K glycoprotein (50 GP); or *f*, mAb VIA41 against the 156K glycoprotein (156 GP). Arrows indicate the positions of dystrophin (DYS), the voltage-sensitive sodium channel (NaCh), and the α_1 and α_2 subunits of the dihydropyridine (DHP) receptor. Arrowheads denote the positions of the molecular weight standards as indicated.

METHODS. Heavy microsomes were prepared from rabbit skeletal muscle²⁵ and washed twice with 0.6 M KCl in 50 mM Tris-HCl (pH 7.4), 0.165 M sucrose, 0.1 mM phenylmethylsulphonyl fluoride and 0.75 mM benzamidine to remove contractile proteins. KCl-washed membranes (1 g) were solubilized in 1.0% digitonin, 0.5 M NaCl, and protease inhibitors as previously described¹². After removal of the ryanodine receptor by immunoaffinity chromatography²⁶, the digitonin-solubilized membranes were circulated overnight on a 40 ml WGA-Sepharose column, washed extensively, then eluted with three column volumes of 0.3 M *N*-acetylglucosamine. Eluted fractions containing dystrophin were applied to a 3-ml DEAE-cellulose column and sequentially eluted with the following NaCl concentrations in buffer A

(0.1% digitonin, 50 mM Tris-HCl, pH 7.4, 0.75 mM benzamidine, 0.1 mM PMSF): 0, 25, 50, 75, 100, 110 and 175 mM. Sucrose gradients (12.5 ml linear 5–20% sucrose) containing 0.5 M NaCl and 0.01% NaN_3 in buffer A were prepared using a Beckman density gradient former. Dystrophin complex, which eluted in fraction 2 (3 ml) from the DEAE-column 175 mM NaCl wash, was concentrated to 0.5 ml in a Centricon-100 (Amicon), layered on a sucrose gradient and overlain with 0.5 ml of buffer A containing 175 mM NaCl and 0.01% NaN_3 . Gradients were centrifuged at 4 °C in a Beckman VTi 65.1 vertical rotor for 90 min at 200,000*g*. Fractions (0.6 ml) were collected from the top of the gradients using an ISCO Model 640 density gradient fractionator. Gradient fractions were separated by SDS-PAGE²⁷ (3–12% gradient gel) and stained with Coomassie blue (300 μl concentrated with a Centricon-100) or transferred to nitrocellulose (75 μl of fractions in *b*, 25 μl in *c* and *d*, and 50 μl in *e* and *f*) and stained with various antibodies. The blot shown in *e* was prepared from a gel run in the absence of reducing agent plus 10 mM *N*-ethylmaleimide. Gel lanes were scanned with a Hoefer GS 300 scanning densitometer and analysed using GS-360 data analysis software. Polyclonal antisera against a chemically synthesized decapeptide representing the C-terminal of dystrophin was raised in New Zealand white rabbits as described²⁸. Hybridomas were obtained from female BALB/c mice which had been immunized with rabbit skeletal muscle membranes and boosted with WGA eluate²⁹.

complex. Antisera from a rabbit immunized with a chemically synthesized decapeptide representing the predicted C-terminal amino-acid sequence of human dystrophin, stained a single high- M_r protein (Fig. 1c). This protein co-migrated with the predominant isoform of dystrophin stained by sheep polyclonal anti-dystrophin antibodies¹³ (not shown). The antisera showed immunofluorescence staining only on the cell periphery (Fig. 2a), which indicates a restricted localization of dystrophin to the sarcolemma of rabbit skeletal muscle.

A library of monoclonal antibodies against muscle proteins eluted from WGA-Sepharose was also screened for reactivity against components of the dystrophin-glycoprotein complex and by indirect immunofluorescence staining of rabbit skeletal muscle. Of six hybridomas which showed immunofluorescence staining only on the sarcolemma, monoclonal antibodies XIXC2 (Fig. 1d) and VIA4₂ (not shown) were found to stain dystrophin on immunoblots. Both dystrophin monoclonal antibodies are IgM subtypes, and recognized both native and denatured dystrophin. Monoclonal antibody XIXC2 also recognizes the minor lower- M_r isoform of dystrophin which co-purifies with the more abundant isoform (Fig. 1d).

Two of the other sarcolemma-specific monoclonal antibodies were specific for components of the dystrophin-glycoprotein complex (Fig. 1e and 1f). The 50K glycoprotein stained with monoclonal antibody IVD3₁ (Fig. 1e), and has been localized exclusively to the sarcolemmal membrane of rabbit skeletal muscle (Fig. 2c). Monoclonal antibody IVD3₁ recognized only the non-reduced form of the 50K glycoprotein and is not highly cross-reactive. Monoclonal antibody VIA4₁ stained the 156K glycoprotein (Fig. 1f) which co-purified with dystrophin. VIA4₁ recognized the denatured form of the 156K glycoprotein and is highly cross-reactive. It also exhibited weak, but specific immunofluorescent staining of the sarcolemmal membrane (Fig. 2d), consistent with its low affinity for the native 156K glycoprotein. In agreement with the immunofluorescence results, a rabbit membrane preparation greatly enriched in sarcolemmal proteins also showed a substantial enrichment in dystrophin, the 156K and 50K glycoproteins (not shown). Immunofluorescence staining for dystrophin, 50K glycoprotein or the 156K

glycoprotein was equally distributed in fast and slow muscle fibres (not shown).

The association of the dystrophin-glycoprotein complex was also assessed by immunoaffinity adsorption. Immunoaffinity beads were prepared with the monoclonal antibodies XIXC2 (anti-dystrophin) and IVD3₁ (anti-50K glycoprotein) and incubated with the partially purified dystrophin-glycoprotein complex. After pelleting the immunoaffinity beads, the supernatants were removed and the beads were washed extensively. The supernatants and washes were pooled (voids), concentrated, and analysed by SDS-polyacrylamide gel electrophoresis and immunoblotting. The voids from the XIXC2 (anti-dystrophin) and the IVD3₁ (anti-50K glycoprotein) immunoaffinity beads contained no dystrophin, 59K triplet, 50K, 43K doublet or 35K proteins as detected by Coomassie blue staining (Fig. 3a). Both the XIXC2 (anti-dystrophin) and IVD3₁ (anti-50K glycoprotein) immunoaffinity beads quantitatively removed dystrophin from the starting material (Fig. 3c). Analysis of the voids for the 156K (Fig. 3d) and 50K (Fig. 3e) glycoproteins revealed that both the XIXC2 and IVD3₁ immunoaffinity beads selectively adsorbed virtually all of each of these glycoproteins from the voids, whereas the voltage-sensitive sodium channel (Fig. 3b) and the α_1 and α_2 subunits of the dihydropyridine receptor (not shown) remained in the voids. As detected by peroxidase-conjugated WGA (not shown), the 43K and 35K glycoproteins were also adsorbed from the voids. Immunoblots of immunoaffinity beads separated on gels indicated that dystrophin, the 156K and 50K glycoproteins were retained by the beads and not selectively proteolysed (not shown). Initial experiments with monoclonal antibody VIA4₁ (anti-156K glycoprotein) have indicated that it has too low an affinity for the native 156K glycoprotein to be successful in this type of experiment.

Analysis of dystrophic muscle

To investigate whether either of the dystrophin-linked 156K or 50K glycoproteins is affected by the absence of dystrophin, immunoblots of skeletal muscle membranes were prepared from control and *mdx* mice and stained with the various antibodies

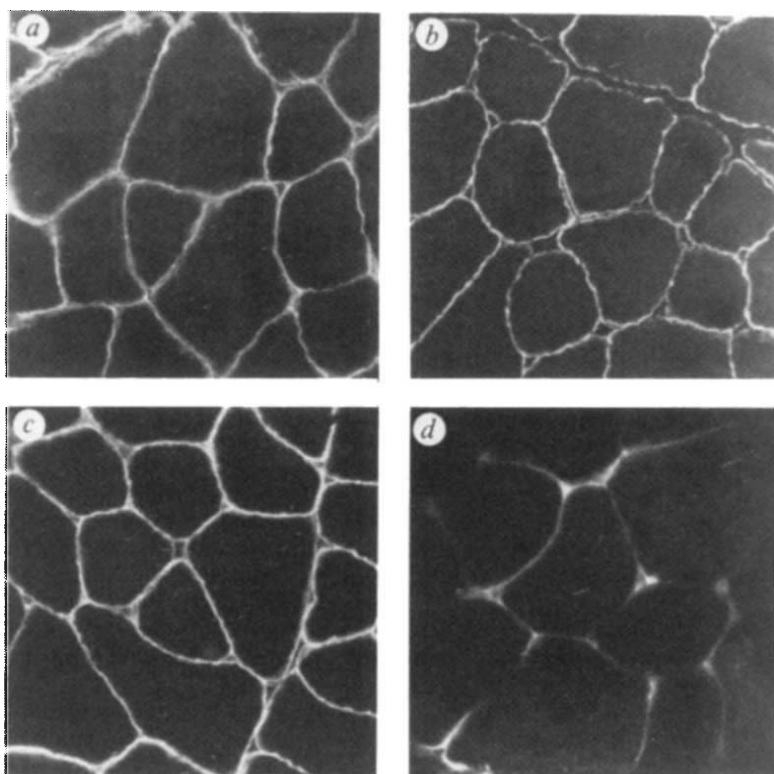
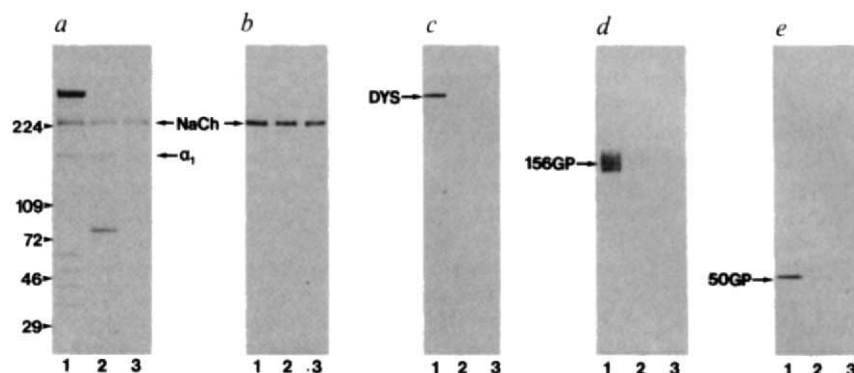


FIG. 2 Immunolocalization of components of the dystrophin complex. Transverse cryostat sections of rabbit skeletal muscle were labelled by indirect immunofluorescence with polyclonal antisera against the C-terminal decapeptide of dystrophin (a), mAb XIXC2 against dystrophin (DYS) (b), mAb IVD3₁ against the 50K glycoprotein (50 GP) (c) and monoclonal antibody VIA4₁ against the 156K glycoprotein (156 GP) (d) (magnification, 250 $\times$). Staining of the cryostat sections was not observed with nonimmune serum, nor was there any nonspecific binding to the tissue by fluorescein-labelled secondary antibody.

METHODS. The indirect immunofluorescence labelling of fixed 8- μ m transverse cryostat sections from rabbit gastrocnemius was carried out as described²⁹. Sections were pre-incubated for 20 min with 5% normal goat antiserum in PBS buffer, followed by a 2 h incubation at 37 °C with the primary antibody (hybridoma supernatants or 1:1,000 diluted antiserum). After washing in PBS, the sections were further incubated for 30 min at 37 °C in PBS with a 1:50 dilution of FITC-labelled goat F(ab')₂ anti-mouse IgG or anti-rabbit IgG and subsequently examined in a Leitz fluorescence microscope.

FIG. 3 Immunoabsorption of the dystrophin-glycoprotein complex. Fraction 2 (125 μ l in a, 25 μ l in b-e) eluted from the 175 mM NaCl wash of the DEAE-cellulose column described in Fig. 1 before treatment (lane 1), the XIXC2 affinity column void (125 μ l in a, 25 μ l in b-e) (lane 2) or the IVD3₁ affinity column void (25 μ l) (lane 3), stained with Coomassie blue (a), mAb G/C6 against the sodium channel (NaCh) (b), polyclonal antisera to the C-terminal decapeptide of dystrophin (DYS) (c), mAb VIA4₁ (156 GP) (d), or mAb IVD3₁ (50 GP) (e). Molecular weight standards (arrowheads) were the same as those used in Fig. 1.

METHODS. Immunoaffinity beads³⁰ were equilibrated with buffer A containing 0.5 M NaCl and then incubated (12 h) with 0.75 ml of fraction 2 from the 175 mM NaCl wash of the DEAE-cellulose column (Fig. 1). After pelleting, the supernatants were decanted (voids) and the affinity beads were washed with 5 aliquots (0.7 ml) of buffer A containing 0.5 M NaCl. The void from each affinity column and the five washes were pooled and concentrated to 375 μ l in a Centricon-100. In addition, 0.75 ml of fraction



2 was diluted to 4.2 ml, concentrated to 375 μ l and used as control. Column voids were analysed by SDS-PAGE and immunoblotting as described in Fig. 1. Monoclonal antibody G/C6 against the skeletal muscle sodium channel³¹ was the gift of Dr Robert Barchi (University of Pennsylvania).

(Fig. 4). Staining with polyclonal antisera against the C-terminal decapeptide of dystrophin revealed that dystrophin was completely absent from *mdx* mouse membranes (Fig. 4a). In addition, comparison of normal and *mdx* mouse with immunostaining by monoclonal antibody VIA4₁ against the 156K glycoprotein revealed that the 156K glycoprotein was absent or greatly reduced in *mdx* mouse membranes (Fig. 4b). Staining of identical transfers with sheep polyclonal antisera against either the ryanodine receptor (Fig. 4c) or the dihydropyridine receptor (Fig. 4d) did not differ between control and *mdx* mouse muscle membranes. Monoclonal antibody IVD3₁ against the 50K glycoprotein did not cross-react with normal mouse membranes and thus could not be evaluated. The absence of the 156K glycoprotein was also confirmed using SDS muscle extracts (not shown) instead of isolated membranes from control and *mdx*

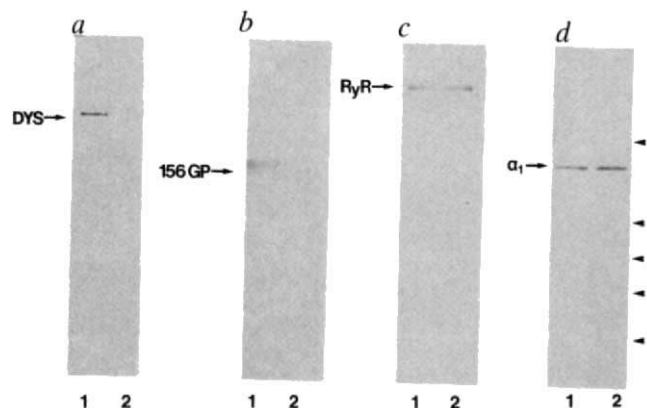


FIG. 4 Immunoblot analysis of control and *mdx* mouse muscle membranes. Immunoblots stained with polyclonal antisera against the C-terminal decapeptide of dystrophin (DYS) (a), mAb VIA4₁ against the 156K glycoprotein (156 GP) (b), sheep polyclonal anti-ryanodine receptor antibody (RyR) (c) and sheep polyclonal anti-DHP receptor antibody (α_1) (d) are shown. Lanes (1) and (2) for each panel consist of equal amounts of muscle membrane protein from control and *mdx* mice, respectively (300 μ g per lane in (a) and (b), 150 μ g per lane in (c) and (d)). Molecular weight standards (arrowheads) were the same as Fig. 1.

METHODS. Membranes from control and *mdx* mice (gifts of Dr Richard Strohmen and Dr Richard Enriqui of University of California, Berkeley and Davis, respectively) were prepared in 10% sucrose, 76.8 mM aprotinin, 0.83 mM benzamidin, 1 mM iodoacetamide, 1.1 μ M leupeptin, 0.7 μ M pepstatin A, 0.23 mM PMSF, 20 mM Tris-maleate, pH 7.0, by centrifuging muscle homogenates for 15 min at 14,000g and subsequently pelleting the supernatant for 30 min at 125,000g, followed by KCl washing as described in Fig. 1. Control and *mdx* mouse muscle membranes were analysed by SDS-PAGE and immunoblotting as described in Fig. 1. The amount of 156K glycoprotein in each preparation was estimated densitometrically from autoradiographs of identical blots incubated with ¹²⁵I-labelled sheep anti-mouse secondary antibody³².

mice. Estimation of the amount of 156K glycoprotein remaining in the *mdx* muscle membranes using ¹²⁵I-labelled secondary antibodies and total membrane preparations from four control and four *mdx* mice revealed an average reduction of 85% in *mdx* muscle.

Total muscle extracts were also prepared from biopsy samples of normal controls and DMD patients (obtained from the Department of Neuropathology, University of Iowa). The dystrophic samples showed no staining with antibodies against dystrophin by indirect immunofluorescence microscopy (not shown) and immunoblotting (Fig. 5a). In contrast to the normal muscle extract, the three DMD samples showed greatly reduced staining for the 156K glycoprotein (Fig. 5b). In contrast, identical immunoblots stained with monoclonal antibodies against the Ca²⁺-dependent ATPase (Fig. 5c) revealed no difference in the staining intensity between normal and dystrophic muscle samples. Again, the amount of 156K glycoprotein was reduced by about 90% in DMD samples.

Discussion

We have presented evidence for the existence of a large oligomeric complex (~18S) containing dystrophin, a 59K triplet

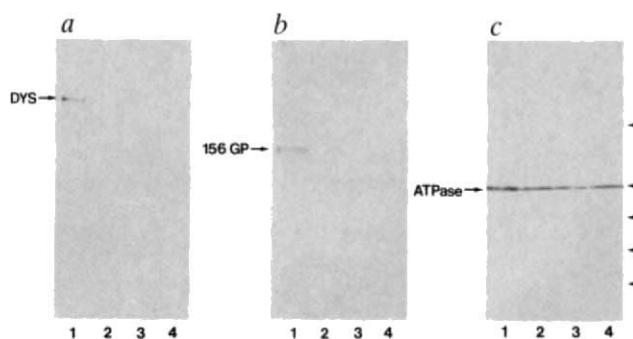


FIG. 5 Immunoblot analysis of normal and dystrophic human muscle biopsies. Immunoblots stained with mAb VIA4₁ against dystrophin (DYS) (a), mAb VIA4₁ against the 156K glycoprotein (156 GP) (b), or mAb IID8 against the (Ca²⁺ + Mg²⁺)-ATPase (ATPase) (c) are shown. Lane (1) consists of normal human muscle extract and lanes (2-4) are dystrophic muscle extracts from three DMD patients. Molecular weight standards (arrowheads) were the same as Fig. 1.

METHODS. Frozen muscle biopsy samples (50 mg) were crushed in liquid nitrogen using a mortar and pestle and then prepared for electrophoresis as described⁶. The pulverized muscle samples were transferred to 10 volumes of SDS-PAGE sample buffer (10% SDS, 2 M sucrose, 4% 2-mercaptoethanol, 0.002% bromophenol blue, 260 mM Tris-HCl, pH 6.8), vortexed, and precipitated material allowed to settle. Aliquots (50 μ l) of the SDS-extracted muscle samples were analysed by SDS-PAGE and immunoblotting as described in Fig. 1 and the amount of 156K glycoprotein was estimated as described in Fig. 4.

and four sarcolemmal glycoproteins (156K, 50K, 43K and 36K). At least one of the proteins in the complex is an integral membrane protein, as 1.0% digitonin was necessary to solubilize the dystrophin-glycoprotein complex. The immunoaffinity experiments demonstrate that the complex is tightly associated. To date, a large number of antibodies specific for extracellular matrix proteins, cytoskeletal proteins, plasma membrane pump, channel and receptor proteins have been screened for reactivity against the dystrophin-glycoprotein complex. None of these antibodies to known proteins has demonstrated cross-reactivity to any component of the complex. The elucidation of primary sequences by recombinant DNA techniques should provide clues to the function of the dystrophin-associated glycoproteins.

The surface of DMD myofibres have been reported to show altered¹⁴ or decreased¹⁵ lectin binding, and a 370K glycoprotein is apparently missing from DMD muscle¹⁶. However, to our knowledge this is the first demonstration of the marked deficiency of a glycoprotein that is closely linked to dystrophin. The substantial reduction of the 156K glycoprotein from muscles of *mdx* mice and DMD patients is analogous to findings in erythrocytes of individuals afflicted with hereditary elliptocytosis¹⁷, in which the absence of the cytoskeletal protein band 4.1^{11,18,19} is accompanied by greatly diminished steady-state levels of glycophorin C^{10,11}. In both diseases, the disruption of the cytoskeleton seems to destabilize the plasma membrane and

associated proteins. As antibody probes to other components of the dystrophin complex become available, it will be interesting to determine if any other proteins are also affected in *mdx* and DMD muscle.

How the absence of dystrophin leads to the clinical manifestation of DMD is an unanswered question. Clearly there could be many steps in the disease process, yet we may have identified the first, which is the loss of a dystrophin-associated glycoprotein due to the absence of dystrophin. For example, muscle from *mdx* mice shows elevated intracellular ionized Ca^{2+} levels and corresponding higher net degradation of muscle proteins²⁰. Loss of a dystrophin-anchored protein with a role in the regulation of intracellular calcium could result in elevated intracellular calcium levels and lead to the reported activation of calcium-dependent protease activities²¹. Such a mechanism could explain the abnormal muscle protein degradation and fibre necrosis of dystrophic muscle.

The absence of dystrophin-associated proteins in dystrophic muscle may complicate the therapeutic efficacy of myoblast transfer²², as the reintroduction of dystrophin synthesis might not lead to recovery of associated protein levels and thus could necessitate treatment at only one particular developmental stage. Finally, a deficiency or defect in a dystrophin-associated glycoprotein could perhaps explain the DMD-like symptoms observed in suspected autosomal-recessive patients^{23,24} that express apparently normal dystrophin. □

Received 22 December 1989; accepted 26 March 1990.

- Hoffman, E. P., Brown, R. H. & Kunkel, L. M. *Cell* **51**, 919-928 (1987).
- Zubrzycka-Gaarn, E. E. et al. *Nature* **333**, 466-469 (1988).
- Arahata, K. et al. *Nature* **333**, 861-866 (1988).
- Bonilla, E. et al. *Cell* **54**, 447-452 (1988).
- Watkins, S. C., Hoffman, E. P., Slayter, H. S. & Kunkel, L. M. *Nature* **333**, 863-866 (1988).
- Hoffman, E. P. et al. *New Engl. J. Med.* **318**, 1363-1368 (1988).
- Cooper, B. J. et al. *Nature* **334**, 154-156 (1988).
- Koenig, M., Monaco, A. P. & Kunkel, L. M. *Cell* **53**, 219-228 (1988).
- Davison, M. D. & Critchley, D. R. *Cell* **52**, 159-160 (1988).
- Alloisio, N. et al. *Biochim. biophys. Acta* **816**, 57-62 (1985).
- Mueller, T. J. & Morrison, M. *Erythrocyte Membrane 2: Recent Clinical and Experimental Advances* 95-112 (Liss, New York, 1981).
- Campbell, K. P. & Kahl, S. D. *Nature* **338**, 259-262 (1989).
- Knudson, C. M., Hoffman, E. P., Kahl, S. D., Kunkel, L. M. & Campbell, K. P. *J. biol. Chem.* **263**, 8480-8484 (1988).
- Bonilla, E., Schotland, D. L. & Wakayama, Y. *Ann. Neurol.* **4**, 117-123 (1978).
- Capaldi, M. J., Dunn, M. J., Sewry, C. A. & Dubowitz, V. *J. neurol. Sci.* **63**, 129-142 (1984).
- Capaldi, M. J., Dunn, M. J., Sewry, C. A. & Dubowitz, V. *J. neurol. Sci.* **68**, 225-231 (1985).
- Palek, J. *Clin. Haematol.* **14**, 45-87 (1985).
- Alloisio, N., Dorleac, E., Giro, R. & Delaunay, J. *Hum. Genet.* **59**, 68-71 (1981).

- Tchernia, G., Mohandas, N. & Shohet, S. B. *J. clin. Invest.* **68**, 454-460 (1981).
- Turner, P. R., Westwood, T., Regen, C. M. & Steinhardt, R. A. *Nature* **335**, 735-738 (1988).
- Nagy, B. & Samaha, F. J. *Ann. Neurol.* **20**, 50-56 (1986).
- Partridge, T. A., Morgan, J. E., Coulton, G. R., Hoffman, E. P. & Kunkel, L. M. *Nature* **337**, 176-179 (1989).
- Francke, U., Darras, B. T., Hersh, J. H., Berg, B. O. & Miller, R. G. *Am. J. hum. Genet.* **45**, 63-72 (1989).
- Zatz, M., Passos-Bueno, M. R. & Rapaport, D. *Am. J. med. Genet.* **32**, 407-410 (1989).
- Mitchell, R. D., Palade, P. & Fleischer, S. *J. Cell Biol.* **95**, 1008-1016 (1983).
- Imagawa, T., Smith, J. S., Coronado, R. & Campbell, K. P. *J. biol. Chem.* **262**, 16636-16643 (1987).
- Laemmli, U. K. *Nature* **227**, 680-685 (1970).
- Strynadka, N. C. J., Redmond, M. J., Parker, J. M. R., Scraba, D. G. & Hodges, R. S. *J. Virol.* **62**, 3474-3483 (1988).
- Jorgensen, A. O. et al. *Cell Motil. Cytoskel.* **9**, 164-174 (1988).
- Campbell, K. P. et al. *J. biol. Chem.* **262**, 6460-6463 (1987).
- Casadei, J. M. & Barchi, R. L. *J. Neurochem.* **48**, 773-778 (1987).
- Sharp, A. H. & Campbell, K. P. *J. biol. Chem.* **264**, 2816-2825 (1989).

ACKNOWLEDGEMENTS. We thank Drs Robert Barchi, Richard Strohm, Richard Entringer, Victor Ionasescu and Michael Hart for gifts of antibodies, mice and human tissues used in this study, and William Horne for helpful discussions and Charles Lovig (University of Iowa Cancer Center). K.P.C. is an Investigator of the Howard Hughes Medical Institute. J.M.E. is an NINCDS postdoctoral fellow. This work was supported by the Muscular Dystrophy Association.

LETTERS TO NATURE

Flat-spectrum radio sources: cosmic conspiracy or relativistic neutrons?

Peter M. Giovanoni & Demosthenes Kazanas

Laboratory for High Energy Astrophysics,
NASA/Goddard Space Flight Center,
Greenbelt, Maryland 20771, USA

RADIO-LOUD active galactic nuclei (AGNs) tend to show extended emission in the form of jets and an unresolved central core, from which the jets presumably originate. The intensity spectrum of the core varies smoothly from $10^{8.5}$ to 10^{16} Hz in frequency¹, and is flat (constant with frequency) between 10^9 and 10^{12} Hz, implying that a single emission mechanism, probably synchrotron radiation, is responsible. Because synchrotron emission below $\sim 10^{13}$ Hz would be self-absorbed for a uniform source under conditions typical of AGNs, inhomogeneous models have been used to account for these spectra. These models, which require the magnetic field

strength and the electron density and energy to vary according to strict power laws, reproduce the flat spectra by a number of self-absorbed components²⁻⁴, but they do not explain the origin of radiating electrons at $\sim 10^{18}$ cm from the AGN core, in view of the very short lifetime of the electrons in an AGN environment. Here we propose that energy is transported from the central source by relativistic neutrons, which travel freely over a large volume and decay into relativistic protons. The protons produce secondary electrons which generate the observed radiation. The photon spectra thus produced are largely model independent and flat.

Relativistic neutrons can result from the collisions of high-energy protons with the ambient gas⁵, or high-energy protons with the ambient photon field if their energies are sufficiently large⁶. Actually, in high-energy pp collisions the leading particle carries most of the energy⁷. In about half of these interactions the leading proton is converted into a neutron, so the neutrons can carry away ~ 25 -50% of the total luminosity available (here we assume that this fraction is 25%). But unlike protons, the relativistic neutrons are not tied to the ambient magnetic field; they readily escape from the system and traverse on the average a distance $r \approx 3 \times 10^{13} \gamma_n$ cm before they decay, γ_n being the neutron Lorentz factor. Therefore, if the neutron distribution

extends to $\gamma_n \approx 10^5$, relativistic neutrons will reach as far as ~ 1 pc from the 'central engine'. The relativistic protons resulting from the neutron decay produce secondary relativistic electrons (and positrons) throughout this region *in situ*, thus successfully addressing the question of the origin of these particles at such large distances from the continuum source. In addition, the emission of radiation over $r \approx 1$ pc is intimately related, through the neutron transport, to the conditions in the central source, thus severely restricting the assumptions that can be made about the population of these relativistic radiating electrons.

We shall assume that at the 'central engine' there exists a steady-state relativistic proton distribution proportional to γ^{-s} , $s \approx 2$, that is, similar to that expected from a strong shock^{8,9}. This steady state is considered to be the result of continuous injection of relativistic protons by the 'engine' and their removal by nuclear collisions¹⁰. As argued above, the proton distribution will give rise to a steady flux of neutrons $F_n(\gamma_n, r) \propto \gamma_n^{-s} r^{-2} \exp(-r/c\tau\gamma_n) \text{ cm}^{-2} \text{ s}^{-1}$, where τ is the rest-frame lifetime of the neutron, $\sim 10^3$ s. The γ_n^{-s} factor reflects the spectrum of the primary protons responsible for the neutron production, the r^{-2} factor the geometric dilution, and the exponential factor their depletion through decay (there is, in general, another depletion due to their interactions with ambient gas but it is energy independent and has been omitted here for simplicity; it will be argued later, that this term must be insignificant to produce a radio-loud object).

The differential spectrum of protons resulting from the neutron decay Q_p , is equal to the number of neutrons decaying in a radius interval dr , or more precisely to the divergence of the neutron flux, $\nabla \cdot \mathbf{F}_n$. For spherical symmetry this is

$$Q_p(\gamma_p, r) = \frac{1}{r^2} \frac{d}{dr} [r^2 F_n(\gamma_n, r)] \\ = \left(\frac{r_0}{r}\right)^2 \frac{I_n}{c\tau} \gamma_p^{-(s+1)} \exp\left(\frac{-r}{c\tau\gamma_p}\right) \text{ cm}^{-3} \text{ s}^{-1} \quad (1)$$

where I_n is the normalization of the neutron flux at the radius r_0 (also bearing in mind that γ_p is almost identical to γ_n). At low energies, the proton injection spectrum has a kinematic cutoff at $\gamma_{p,b} \approx r/c\tau$ due to the decay of the lower-energy neutrons, which increases linearly with r . Above this cutoff the spectrum is steeper than the neutron spectrum by one power of γ , due to fewer decays of high-energy neutrons relative to

low-energy neutrons in a shell of thickness dr . This extra steepening is important for obtaining a 'flat' radio spectrum.

The decay-produced relativistic protons couple to the ambient magnetic field and incorporate themselves into the local plasma which for simplicity is considered to be at rest (see, however, ref. 11). In a steady state they will deposit all their energy through nuclear collisions in the form of relativistic e^-e^+ ($\sim 20\%$), π^0 -decay γ -rays ($\sim 30\%$) and neutrinos ($\sim 50\%$). The Lorentz factor of the resulting electrons is $\gamma_e \approx 60\gamma_p$, and the energies of the γ -rays are $E_\gamma \approx 70\gamma_p \text{ MeV}$.

The resulting injection rate for secondary relativistic electrons, $Q_e(\gamma_e, r)$, is very similar to that of protons except that the kinematic cutoff is at a Lorentz factor $\gamma_{e,b} \approx 60\gamma_{p,b}$; its normalization is also different, determined by demanding that the secondary electrons, of $\gamma_e \approx 60\gamma_p$, carry $\sim 20\%$ of the proton luminosity. $Q_e(\gamma_e, r)$ is given approximately by

$$Q_e(\gamma_e, r) \approx 2.14 \times 10^4 \left(\frac{r_0}{r}\right)^2 \frac{I_n}{c\tau} \gamma_e^{-(s+1)} \\ \times \exp\left(\frac{-60r}{c\tau\gamma_e}\right) \text{ cm}^{-3} \text{ s}^{-1} \quad (2)$$

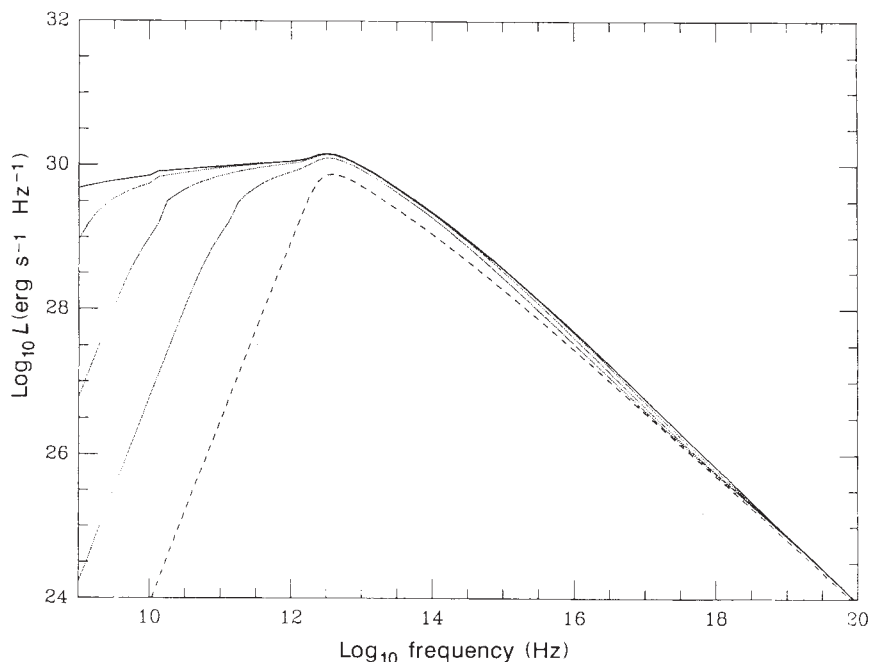
The calculation of the steady-state emission requires the calculation of the electron steady-state distribution as a function of radius, $N_e(\gamma_e, r)$, given the injection spectrum of equation (2). Earlier treatments of the problem²⁻⁴ assumed this distribution, thereby bypassing the question of the origin of the relativistic electrons. It was nonetheless shown (see, for example, ref. 3) that, assuming the power-law forms $N_e(\gamma_e, r) \propto r^{-n} \gamma_e^{-p}$ and $B(r) \propto r^{-m}$, the superposition of synchrotron self-absorbed components gives a spectrum $F_\nu \propto \nu^{-\alpha_i}$ with

$$\alpha_i = \frac{3m + 5n + 2p(m-1) - 13}{[2n + m(p+2) - 2]} \quad (3)$$

The exponent α_i can be set to any value (in particular zero) by the appropriate choice of m , n and p . The fact that in a class of objects (the flat-spectrum radio sources) the combination of m , n and p is such that $\alpha \approx 0$ came to be known as 'cosmic conspiracy' (see, for example, ref. 12).

In the present model, the free independent choice of these parameters is not allowed because they are related to the dynamics of energy release near the black hole through the energy transport by neutrons and the resulting secondary-elect-

FIG. 1 Logarithmic plot of L_ν against ν for a source with luminosity in neutrons (at $r=r_0$) $L_n = 10^{45} \text{ erg s}^{-1}$, magnetic field at $r \leq r_0$, $B_0 = 300 \text{ G}$, central source size $r_0 = 3 \times 10^{14} \text{ cm}$ and $\gamma_{p, \text{max}} = 10^5$. The solid line is the final spectrum, that is, the spectrum at infinity. The dashed line shows the spectrum from just the central source, that is, $r \leq r_0$. The dotted lines show the spectrum of the radiation produced inside $r = 10r_0$, $10^2 r_0$ and $10^3 r_0$.



tron production. More importantly, these parameters can be calculated once the electron injection rate as a function of energy and radius, $Q_e(\gamma_e, r)$, is given. The resulting steady-state electron distribution function is then

$$N_e(\gamma_e, r) = \frac{1}{\dot{\gamma}} \int_{\gamma_e}^{\infty} Q_e(\gamma'_e, r) d\gamma'_e \quad (4)$$

where $\dot{\gamma}$ is the electron energy loss rate. In the vicinity of an AGN the dominant energy loss processes for relativistic electrons are synchrotron and inverse Compton scattering. For these processes the energy loss rate (in the Thomson approximation) is

$$\dot{\gamma}(r) = \frac{4}{3} \frac{\sigma_T \gamma_e^2}{m_e c} \rho(r) \quad (5)$$

where $\rho(r) = L/4\pi r^2 c + B^2(r)/8\pi$ is the total energy density in photons and magnetic fields. If the magnetic field is in equipartition with the photon energy density (or tied to an outflowing wind), $B(r) \propto 1/r$ and hence $\dot{\gamma}(r) \propto 1/r^2$ for $r \geq r_0$, whereas it is assumed to be constant for $r < r_0$.

Under these conditions, equation (4) indicates that for $\gamma > \gamma_{e,b} \approx 60r/c\tau$, the electron distribution function is $N_e(\gamma_e) = A_{>} \gamma_e^{-q}$, $q = s + 2$, with the normalization $A_{>}$ independent of the distance r because the geometric dilution factor r^{-2} in $Q_e(\gamma_e, r)$ is cancelled by the corresponding smaller $\dot{\gamma}$ ($\dot{\gamma} \propto r^{-2}$). For $\gamma_e < \gamma_{e,b}$ the electron spectrum is a power law of index 2, $N_e(\gamma_e) = A_{<} \gamma_e^{-2}$, independent of the details of injection, the factor γ_e^{-2} simply reflecting the energy dependence of the electron loss rate (see, for example, ref. 13), but with a normalization $A_{<}$ that is r -dependent. The requirement then that these two spectra join at the break energy $\gamma_{e,b}$ (with $\gamma_{e,b} \propto r$) forces the normalization of $N_e(\gamma_e, r)$ for $\gamma < \gamma_{e,b}$ to be $A_{<} \propto \gamma_{e,b}^{(q-2)} = \gamma_{e,b}^s \propto r^{-2}$ for $s \approx 2$. It is thus important to note that the radial dependence of $A_{<}$ is closely tied to the spectral index q at $\gamma > \gamma_{e,b}$ through the matching of the two spectra at $\gamma = \gamma_{e,b}$; the $A_{<} \propto r^{-2}$ dependence is in fact obtained only if $q \approx 4$ ($s \approx 2$).

The assumption of energy transport by neutrons therefore forces a relativistic-electron distribution that is a power law of index $p = 2$, with normalization $A_{<} \propto 1/r^2$ (that is, $n = 2$), and as argued $B(r) \propto 1/r$ ($m = 1$) is a reasonable assumption for the magnetic-field radial dependence. This is precisely one of the sets of values of m , n and p that results in $\alpha_t = 0$, thereby producing flat radio spectra in the $\sim 10^9$ – $10^{12.5}$ Hz frequency range. (The emission by electrons of energies $\gamma \approx \gamma_{e,b}$ is always optically thin and hence this part of the electron distribution should not be used in equation (3).)

We have performed in detail the calculations qualitatively outlined above, taking into account the exact form of the electron losses and emissivities (employing the Klein-Nishina cross-sections), the full treatment of the synchrotron self-Compton interactions (including iterations until convergence when the inverse-Compton luminosity was higher than that of synchrotron)¹⁴, the precise calculation of the synchrotron emission and absorption coefficients, and the reinjection of e^+e^- pairs as a result of photon-photon absorption of the π^0 -decay γ -rays. The detailed calculations have essentially confirmed our qualitative analysis. A sample of such a calculation is given in Fig. 1 where the spectrum of a source with luminosity $L = 10^{45}$ erg s⁻¹, size $r_0 = 3 \times 10^{14}$ cm and $B(r_0) = 300$ G is presented, spanning the range 10^9 – 10^{20} Hz.

Our model also provides for radio-quiet objects with the introduction of only one more parameter: the density of the ambient accreting matter (or photons) in the vicinity of the 'central engine', the effect of which is to convert the escaping neutrons into protons by means other than their decay, namely, by collisions. Interestingly, the cross-section for these processes increases logarithmically with energy, resulting in shorter mean free paths for the higher-energy neutrons. Thus, when the ambient density becomes larger than a certain value, these collisions preferentially deplete the highest-energy neutrons

which are responsible for providing the relativistic electrons at the largest radii and hence at the lowest frequencies; as a result the emission at $\nu \leq 10^{12.5}$ Hz is suppressed and the final spectrum is similar to that of a radio-quiet object. Because the neutron mean free path scales like $\dot{m}/\dot{m}_E$ ($\dot{m}_E$ is the Eddington accretion rate) for a simple spherical-accretion model¹⁰, the above arguments indicate that the radio-loud objects should operate at lower values of this parameter than the radio-quiet ones.

The present treatment involves a number of simplifying assumptions that might, at first sight, be objected to on observational grounds. One of them is that of spherical symmetry, although similar spherically symmetric models do exist in the literature²⁻⁴. Observations do indicate the presence of an unresolved core and linear structures composed of individual components (blobs) of angular sizes of a few milliarcseconds, resolved by very-long-baseline interferometry, that seem to be aligned with the larger scale (arcsec) radio jets. We point out, however, that our models pertain mainly to the unresolved core rather than to these individual 'blobs', which should represent emission from much larger distances. The unresolved core seems to contain most of the flux at frequencies ≥ 5 GHz and at the same time it is in itself 'flat' down to 5–10 GHz. For example, in the case of NRAO140, extensively analysed in ref. 15, these individual components appear at linear scales ~ 10 –50 pc, that is, more than an order of magnitude larger than those of our models. The spherical model outlined above should in any case be an excellent approximation to a conical jet. Considering the rather isotropic emission of neutrons at the central source, these jets would be expected to have large opening angles (≥ 1 sr), at least at scales ≤ 1 pc. This is not as detrimental a feature as it might seem considering the collimated jets observed at larger (arcsec) angular scales, because collimation may in fact be affected by the pressure gradients of the underlying galaxy¹⁶. Large-opening-angle jets over length scales ≤ 1 pc, which collimate at larger scales ($\gg 1$ pc) may also help resolve the problem of the statistics of unbeamed sources¹². The apparently superluminally moving 'blobs' could then represent emission from only a small fraction of this 'large angle' jet, namely the fraction that lies in a small angle $\theta \approx \Gamma_{\text{jet}}^{-1}$, (Γ_{jet} being the jet's bulk Lorentz factor) to the observer's line of sight.

Another simplification pertaining to the observed spectra, the main thrust of the model, is the assumption that the emitting plasma is static. We have also run models relaxing this assumption by including adiabatic losses in the calculations (adding the term $3\dot{\gamma}/r \approx 3c/r$ in $\dot{\gamma}$). In certain cases this modified the emission, yielding slightly inverted spectra ($\alpha_t \approx -0.3$) at the lowest frequencies ($\sim 10^9$ Hz), which are not in disagreement with the data¹. A related issue is that of source variability (irrelevant in the present static, steady-state models), which pertains to the proton collision timescale. Simple spherical-accretion scaling indicates densities $\sim 10^4$ – 10^5 cm⁻³, at distances ≤ 1 pc appropriate for the lowest frequency (GHz) emission, which yield timescales 30–300 yr. The lower figure is in fact consistent with the observed variability of NGC1275 (ref. 17). Faster variability is expected at higher frequencies ($t_{\text{var}} \propto \nu^{-3/2}$), as observed. We will address this important issue in detail in a future publication. Finally, we have also neglected the effects of photon-hadron interactions, which may become important for neutrons of Lorentz factors $\gamma_n > 10^5$, as indicated in ref. 18.

Besides the inhomogeneous spherical models²⁻⁴, emission from jet models for these objects also exist¹⁹; we argue, however, in favour of the spherically symmetric (or conical jet) ones on the basis of their simplicity and the smaller number of free parameters they involve. Even in this more restricted class of models, the present one is significantly more constrained in the allowed values of m , n and p . In particular, the value of p is completely model-independent¹³, as it simply reflects the form of the energy losses of relativistic electrons. The value of n does depend on the assumed proton spectrum in the central source ($N_p \propto \gamma_p^{-2}$ being a reasonable assumption); however, the value

of n required to produce a flat radio spectrum, namely $n = 2$, is obtained only after this proton spectrum has been supplemented by the specifics of the energy transport by neutron escape and subsequent decay (such as the steepening of the proton spectra to γ^{-3}). Finally the choice $B(r) \propto 1/r$ is the most natural one by the arguments already given. It could therefore be argued that the flat spectra of the compact radio sources are an issue not of 'cosmic conspiracy', but rather of 'cosmic necessity', in the sense that their form can result from the simplest, most conservative assumptions, once the presence of relativistic protons in the 'central engine' of these sources is accepted. $\square$

Received 20 December 1989; accepted 29 March 1990.

- Landau, R. *et al. Astrophys. J.* **308**, 78–92 (1986).
- Condon, J. J. & Dressel, L. L. *Astrophys. Lett.* **15**, 203–207 (1973).
- Marscher, A. P. *Astrophys. J.* **216**, 244–256 (1977).
- Band, D. L. & Grindlay, J. E. *Astrophys. J.* **308**, 576–591 (1985).
- Kazanas, D. & Ellison, D. C. *Nature* **319**, 380–382 (1986).
- Kirk, J. G. & Mastichiadis, A. *Astr. Astrophys.* **213**, 75–79 (1989).
- De Kerret, H. *et al. Phys. Lett.* **B63**, 477–482 (1976).
- Bell, A. R. *Mon. Not. R. astr. Soc.* **182**, 147–156 (1978).
- Blandford, R. D. & Ostriker, J. P. *Astrophys. J.* **221**, L29–L32 (1978).
- Kazanas, D. & Ellison, D. C. *Astrophys. J.* **304**, 178–187 (1986).
- Sikora, M. *et al. Astrophys. J.* **341**, L33–L36 (1989).
- Phinney, E. S. in *Astrophysics of Active Galaxies and Quasi-Stellar Objects* (ed. Miller, I. S.) 453–496 (University Science Books, Mill Valley, 1985).
- Kazanas, D. *Astrophys. J.* **287**, 112–115 (1984).
- Zdziarski, A. A. & Lamb, D. Q. *Astrophys. J.* **309**, L79–L82 (1985).
- Marscher, A. P. *Astrophys. J.* **334**, 552–559 (1988).
- Krautter, A., Henriksen, R. N. & Lake, K. *Astrophys. J.* **269**, 81–93 (1983).
- Backer, D. C. in *Superluminal Radio Sources* (eds Zensus, J. A. & Pearson, T. J.) 76–82 (Cambridge University, 1987).
- Sikora, M. *et al. Astrophys. J.* **320**, L81–L85 (1987).
- Königl, A. *Astrophys. J.* **243**, 700–709 (1981).

ACKNOWLEDGEMENTS. We thank A. Dar, A. Marscher and R. Streitmatter for discussions.

Variation of low-order acoustic solar oscillations over the solar cycle

Y. Elsworth*, R. Howe*, G. R. Isaak*,
C. P. McLeod* & R. New†

* School of Physics and Space Research, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

† Department of Applied Physics, Sheffield Polytechnic, Sheffield S1 1WB, UK

GLOBAL acoustic oscillation modes of the Sun were discovered eleven years ago¹. The possibility of temporal variations in the oscillation frequencies was suggested by fluctuations in the flux of solar neutrinos², and would also be implied by changes in the size of the solar cavity or in the speed of sound within the Sun. Our group has studied solar oscillations for many years^{1,3}; over the past decade, a network of three stations deployed at sites that can permit 24-h data collection has provided p-mode (acoustic) spectra of very high quality. Here we present evidence that the frequencies of the lowest-order ($l \leq 2$) modes have varied over the period of observation (1977–88) in a manner that is correlated with solar activity (as measured by sunspot number). The frequency variation has a peak-to-peak amplitude of $0.46 \pm 0.06 \mu\text{Hz}$, and could reflect variations in the solar dimensions or in the sound speed in the Sun, which might in turn be due to changes in solar temperature and/or magnetic field.

Detections of changes in the Sun's diameter have been reported^{4,5} but they are considered marginal⁶. Luminosity changes⁷ seem to have been established only recently⁸, but these may be a surface phenomenon rather than necessarily implying changes in the eigenfrequencies. There have been several reports on changes in the frequencies of the low- l modes^{9–18}, but some confusing conflicts have arisen (compare, for example, ref. 11 with refs 12 and 13).

Here we present some of the first results from a network of three stations situated around the world in such a way that in ideal conditions 24-h data collection is possible. The data reported cover the years 1977–88 and, in more recent years with the increased coverage resulting from multi-station operation, give rise to p-mode spectra of very high quality. We believe that the combination of long, continuous data records, better signal-to-noise ratio and a better window function provided by multi-station operation provide data of the highest statistical accuracy combined with the smallest systematic errors available at present.

The Birmingham group began measurements of the globally averaged velocity of the solar surface in 1974 (ref. 3). Every year since 1975 (except 1979) measurements have been made in Tenerife in collaboration with the Instituto de Astrofísica de Canarias, with increasing coverage through the years. Since 1984, data acquisition has been almost continuous. The instrument used is a resonant scattering spectrometer¹⁹ which uses a reference cell containing potassium vapour as a standard, by which shifts in the potassium Fraunhofer line at 769.9 nm may be measured. Variations in the observer's velocity resulting from the Earth's spin and orbit provide a continuous calibration of velocity. Timing is provided to a precision of the order of one second and therefore contributes fractional errors of the order of 10^{-7} for each data set. A second station using an updated instrument operating on principles similar to that in Tenerife was deployed in 1981 at Haleakala in Hawaii and was used during the summers until 1986, when a degree of automation

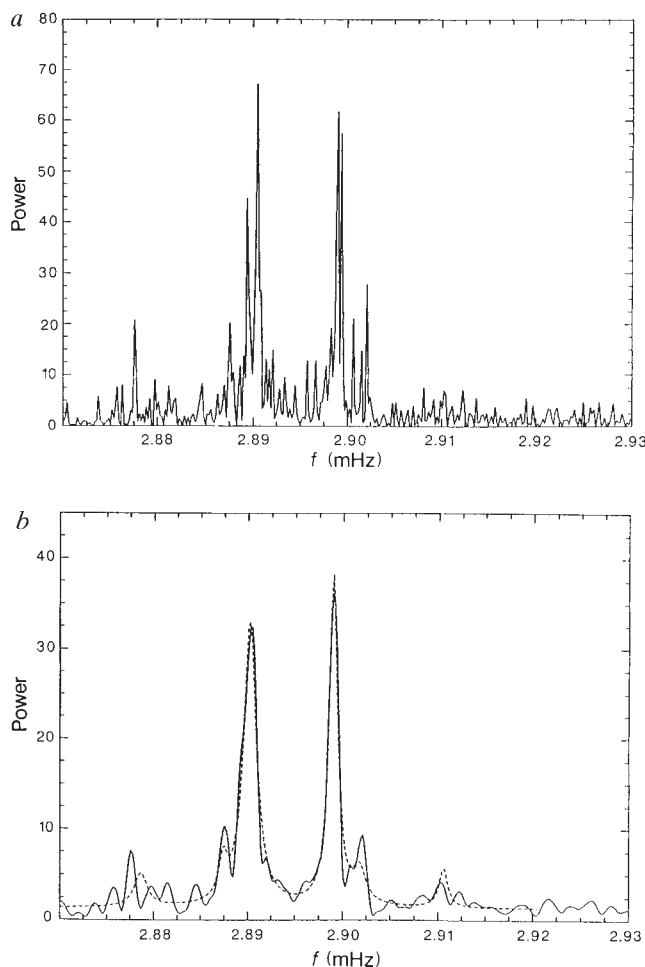


FIG. 1 Portion of a typical power spectrum, in *a*, unsmoothed and *b*, smoothed forms, and fitted by a least-squares technique to a function consisting of two lorentzian peaks of independent height and width with associated sidelobes $\pm 11.57 \mu\text{Hz}$ away on a flat background.

TABLE 1 Composition of power spectra

Year	Julian dates 2440000+	Station	% Time
1977	3336-3379	T	17
1978	3720-3731	T	28
	3745-3760	T	26
1980	4438-4468	T	32
1981	4761-4824	HT	54
1982	5134-5192	HT	48
1983	5507-5569	HT	48
1984	5870-5932	HT	48
1985	6153-6214	T	18
	6215-6276	T	19
1986	6598-6640	HT	23
	6690-6752	TA	23
1987	6978-7040	HTA	44
1988	7237-7299	HTA	33
1988	7350-7412	HTA	61
1988	7431-7492	HTA	30

H, Haleakala, Hawaiian Islands; T, Izana, Tenerife; A, Carnarvon, Western Australia; data were combined as indicated. % time is the percentage of the total observing time in the given interval.

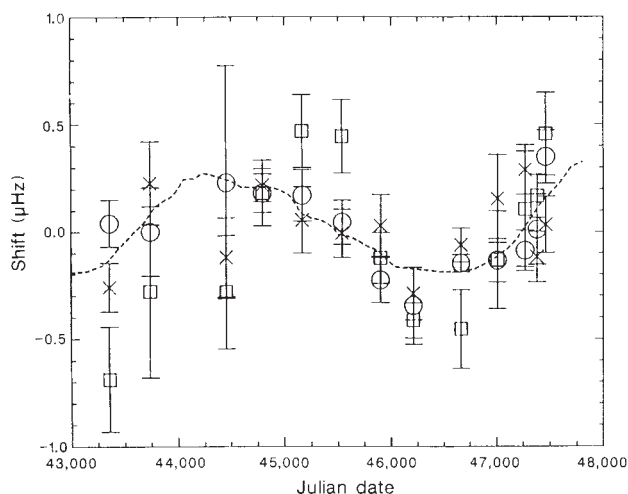


FIG. 2 Variation between 1977 and 1988 of the mean shift for modes with $l=0$, $n=17-24$ (circles), $l=1$, $n=17-24$ (crosses) and $l=2$, $n=16-23$ (squares). The dashed line corresponds to the best linear fit of the shifts to the sunspot number, smoothed over 3 months (Fig. 3).

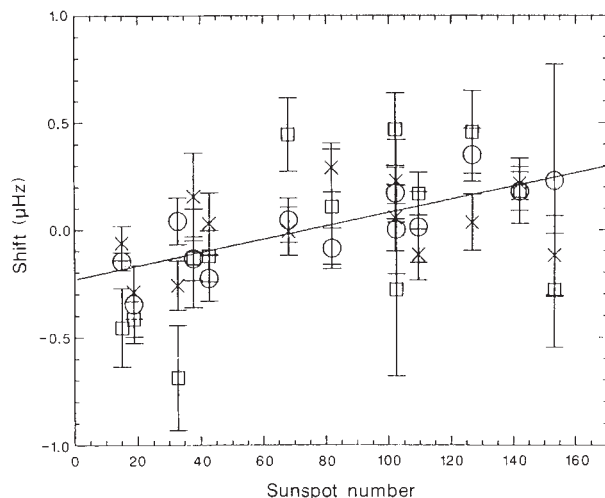


FIG. 3 Correlation between 1977 and 1988 of the mean shift of modes for $l=0$, 1 and 2 with the smoothed sunspot number²³.

allowed more extensive data taking. A third station, incorporating further levels of automation, was deployed in August 1985 at Carnarvon, Western Australia. Despite instrument failures and logistical problems it has proved possible to collect reasonably well filled data stretches covering two-month periods for several years²⁰. These data form the basis for the analysis presented below and for the work on linewidth discussed elsewhere²¹.

Velocity residuals with 42-s spacing, were formed from the raw data by removing the daily trend and substituting zeros for missing points. Power spectra were obtained using a standard fast-Fourier-transform routine. The composition of the spectra used is summarized in Table 1. It should be noted that the signal-to-noise and fill factor were poorer in the early years.

The frequencies of modes were obtained by least-squares fitting to the power spectra, as illustrated for example in Fig. 1. The sidelobes introduced by the day-night cycle, at 11.57 μHz above and below each genuine peak, were taken into account in the fitting function. These sidelobes are very nearly eliminated in good three-station spectra. Tests were carried out for good-quality one-station data which demonstrate that even with the aliases, similar results (but with larger errors) are obtained. Because an unsmoothed power spectrum of modes with lifetimes shorter than, but comparable to, the observation time has a ragged 'spiky' appearance which makes it difficult to decide on the true frequency of the modes, the spectra were first smoothed. A non-distorting function²² was used, with a full width at half maximum of $n/3$ frequency bins for smoothing over n bins: smoothing was performed over several bins increasing from 5 to 49 over the frequency range 2–3.8 mHz, the degree of smoothing being chosen so as to remove the spurious internal structure of the peaks without seriously increasing the measured linewidth. The applicability of the fitting technique is limited by the decreasing linewidth below about 2.2 mHz: above about 3.7 mHz the modes are poorly resolved because of the increased linewidth.

A selected set of strong modes ($n=17-24$) was used in the analysis. The mean frequency of each mode over all the spectra was calculated, allowing the shift of each mode from its mean to be evaluated for each spectrum. The mean shift over all modes of a given l for a given year is thence obtained; those for $l=0$, 1 and 2 in the central region of the spectrum are plotted in Fig. 2 (error bars are standard deviation of the mean). The dashed line corresponds to the best linear fit of shift to sunspot number²³, R_z , smoothed over three months, using the correlation shown in Fig. 3. The correlation coefficient for the 39 points is 0.69.

The acoustic eigenfrequencies considered here correspond to $n \approx 20$ and $l=0$, 1 and 2. If we restrict the discussion to the radial modes we can consider them to be the result of plane standing waves trapped in a cavity of radius $2R_0$ and propagating with a velocity $v(R)$, the local value of which is given approximately by $v = \sqrt{(\gamma kT/m)}$ in the absence of strong magnetic fields (where γ is the adiabatic index and m is the mean molecular weight). The total travel time t is given by

$$t = 2 \int_0^{R_0} \frac{dr}{v(R)}$$

and is strongly weighted towards values at the surface: $v(R_0) = 7 \text{ km s}^{-1}$ compared with $v(0) \approx 500 \text{ km s}^{-1}$. The travel time t is roughly equal to the inverse spacing of the spectrum of eigenfrequencies¹ (135.6 μHz), that is, $t \approx 7,370 \text{ s}$.

The observed peak-to-peak variation of the frequencies, $0.46 \pm 0.06 \mu\text{Hz}$, corresponds to a change in the total travel time of $\sim 1 \text{ s}$. The travel time reaches a maximum at the time of the minimum solar activity. This change could be the result of a change in cavity dimension and/or the velocity of sound in the interior. One of the possible first-order contributions to the variation in sound velocity is variations in temperature T . Sufficiently large changes in γ and m , and in second-order propagation delays due to turbulent motions²⁴, are improbable.

Changes in the strong internal solar magnetic fields, such as have been postulated²⁵ and considered in the context of the eigenfrequencies of the Sun²⁶⁻²⁹, are also possible candidates. □

Received 11 December 1989; accepted 29 March 1990.

1. Claverie, A., Isaak, G. R., McLeod, C. P., van der Raay, H. B. & Roca Cortes, T. *Nature* **282**, 591-594 (1979).
2. Sakurai, K. *Solar Phys.* **74**, 35-41 (1981).
3. Brookes, J. R., Isaak, G. R. & van der Raay, H. B. *Nature* **259**, 92-95 (1976).
4. Gilliland, R. L. *Astrophys. J.* **248**, 1144-1155 (1981).
5. Delache, P., Laclaire, F. & Sadsaoud, H. *Nature* **317**, 416-418 (1985).
6. JOSO Working Group 5 Meeting, November 1989, Rome.
7. Gough, D. O. *Nature* **288**, 639-640 (1980).
8. Willson, R. C. & Hudson, H. S. *Nature* **332**, 810-813 (1988).
9. van der Raay, H. B. in *Theoretical Problems in Stellar Stability and Oscillations* (eds Noels, A. & Gabriel, M.) 215-219 (Université de Liege, Cointe-Ougree, 1984).
10. Isaak, G. R. *et al. Adv. Helio-Asteroseismology IAU Symp.* **123**, 201-204 (1988).
11. Jefferies, S. M., Palle, P. L., van der Raay, H. B., Regulo, C. & Roca Cortes, T. *Nature* **333**, 646-649 (1988).
12. Palle, P. L., Regulo, C. & Roca Cortes, T. *Seismology of the Sun and Sun-like Stars: ESA SP-286*, 285-289 (European Space Agency, Paris, 1988).
13. Palle, P. L., Regulo, C. & Roca Cortes, T. *Astr. Astrophys.* **224**, 253-258 (1989).
14. Woodard, M. F. & Noyes, R. W. *Nature* **318**, 449-450 (1985).
15. Woodard, M. F. & Noyes, R. W. *Seismology of the Sun and the Distant Stars* (ed. Gough, D. O.) 303-304 (Reidel, Dordrecht, 1986).
16. Woodard, M. F. & Noyes, R. W. *Adv. Helio-Asteroseismology IAU Symp.* **123**, 197-200 (1988).
17. Fossat, E., Gelly, B., Grec, G. & Pomerantz, M. A. *Astr. Astrophys.* **177**, L47-L48 (1987).
18. Gelly, B., Fossat, E. & Grec, G. *Astr. Astrophys.* **200**, L29-L31 (1988).
19. Brookes, J. R., Isaak, G. R. & van der Raay, H. B. *Mon. Not. R. astr. Soc.* **185**, 1-17 (1978).
20. Aindow, A. *et al. Seismology of the Sun and Sun-like Stars: ESA SP-286*, 157-160 (European Space Agency, Paris, 1988).
21. Elsworth, Y. P. *et al. Mon. Not. astr. Soc.* **242**, 135-140 (1990).
22. Pollard, J. H. *A Handbook of Numerical and Statistical Techniques* 126 (Cambridge University Press, 1977).
23. *Solar Geophysical Data*, July 1989: No. 539, part I (NOAA, Boulder, Colorado, 1989).
24. Brown, T. M. *Science* **226**, 687-689 (1984).
25. Dicke, R. H. *Solar Phys.* **78**, 3-16 (1982).
26. Isaak, G. R. *Nature* **296**, 130-131 (1982).
27. Dicke, R. H. *Nature* **300**, 693-697 (1982).
28. Gough, D. O. *Nature* **298**, 350-354 (1982).
29. Roberts, B. & Campbell, W. R. *Nature* **323**, 603-605 (1986).

ACKNOWLEDGEMENTS. The assistance of all members, past and present, of the Birmingham and IAC solar oscillations groups is gratefully acknowledged. This work was funded by SERC and CAICYT (Spain).

Superconductivity in the layered compound Li_xNbO_2

Margret J. Geselbracht*, Thomas J. Richardson†
& Angelica M. Stacy*†‡

* Department of Chemistry, University of California, Berkeley, Berkeley, California 94720, USA

† Materials and Molecular Sciences Division, Lawrence Berkeley Laboratory, Berkeley, California 94720, USA

SEVERAL classes of copper oxide compounds are high-temperature superconductors¹⁻³; the highest known transition temperature (T_c) is 122 K for $\text{Ti}_2\text{Ba}_2\text{Ca}_2\text{Cu}_3\text{O}_{11}$ (ref. 4). Superconductivity has also been observed in oxides containing early transition metals, but the superconducting transition temperatures are substantially lower. (The highest known T_c for an early transition metal oxide is 13.7 K, for the spinel compound LiTi_2O_4 (ref. 5).) Whereas all of the copper oxide superconductors have very anisotropic structures, the superconducting oxides of early transition metals discovered up to now have three-dimensional structures. Here we report the discovery of a new class of superconductors, Li_xNbO_2 , with layered structures. At low applied magnetic fields, the magnetic susceptibility greatly decreases to diamagnetic values below 5.5 K for $\text{Li}_{0.45}\text{NbO}_2$ and below 5.0 K for $\text{Li}_{0.50}\text{NbO}_2$; this transition indicates the onset of superconductivity. In this first example of superconductivity in a layered early transition metal oxide, it is interesting to note that the layering has not increased T_c to new levels for early transition metal oxides.

The reduced lithium niobate, LiNbO_2 , has a layered structure analogous to that of MoS_2 . As shown in Fig. 1, the lithium ions occupy octahedral holes between sheets of edge-sharing NbO_6 trigonal prisms⁶. The only study of the deintercalation chemistry of LiNbO_2 is by Kumada *et al.*⁷ who recently reported preparation of Li_xNbO_2 ($0.51 < x < 0.93$) by chemical and electrochemical means. For all materials prepared, they observed temperature-independent paramagnetism between 80-300 K. Owing to the trigonal prismatic coordination of the niobium atoms in LiNbO_2 , the conduction band, which is derived mostly from d_{z^2} orbitals, is narrow⁸. Because this band is completely filled for stoichiometric LiNbO_2 (Nb(III) , d^2), this compound is expected to be semiconducting. Partial removal of lithium atoms results in oxidation of the niobium atoms and an increase in the density of states at the Fermi level as a result of the loss of electrons from the d_{z^2} band. Kumada *et al.* observed an increase in the magnetic susceptibility with decreasing x , consistent with an increase in the density of states. Removal of all the lithium atoms would lead to a half-filled d_{z^2} band (Nb(IV) , d^1), and a high density of states at the Fermi level. BCS-type superconductivity is often observed when the density of states at the Fermi level is high⁹. Although we do not observe superconductivity in LiNbO_2 down to 2 K, partial oxidation of this material by removal of lithium atoms leads to the onset of superconductivity.

The burgundy-red powder, LiNbO_2 , was prepared by the reaction of Li_3NbO_4 and NbO in evacuated fused-silica tubes at 1,050 °C, as in ref. 7. Pressed pellets of the reaction mixture were placed in alumina boats before sealing to minimize reaction with the quartz. Analysis of the powder X-ray diffraction pattern revealed that the product was predominantly LiNbO_2 with only trace amounts of the impurity LiNbO_3 . Hexagonal lattice parameters $a = 2.9063$ (6) Å and $c = 10.447$ (2) Å were derived from the powder diffraction data using silicon as an internal standard. The lithium content was determined by atomic absorption analyses. The product was slightly non-stoichiometric in lithium, $\text{Li}_{0.93}\text{NbO}_2$. Partial oxidation of $\text{Li}_{0.93}\text{NbO}_2$ was carried out by two methods: reaction with bromine and reaction with nitrosyl

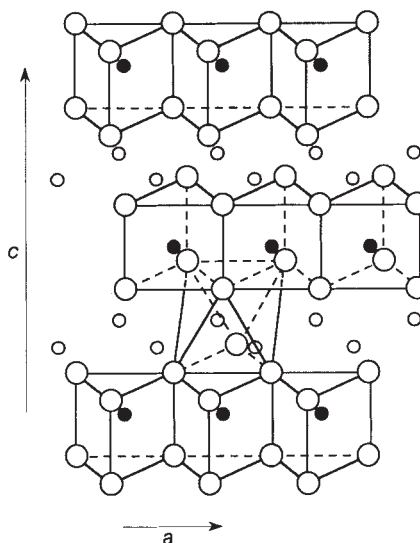


FIG. 1 Schematic of the structure of LiNbO_2 . The layered structure consists of a hexagonal array of niobium atoms (solid circles) sandwiched between arrays of hexagonally ordered oxygen atoms (large open circles). The coordination about the niobium atoms is trigonal prismatic. The lithium atoms (small open circles) are in octahedral sites between the oxygen-niobium-oxygen layers. One row of niobium atoms and two rows of lithium atoms are shown in each of their respective layers. Sufficient oxygen atoms are shown to indicate the coordination spheres.

‡ To whom correspondence should be addressed.

tetrafluoroborate. $\text{Li}_{0.50}\text{NbO}_2$ was prepared by heating to reflux 1.7 g of $\text{Li}_{0.93}\text{NbO}_2$ powder in 10 ml of an acetonitrile solution that was 0.3M in Br_2 for 3 days. $\text{Li}_{0.45}\text{NbO}_2$ was prepared by stirring a 70-ml slurry of 0.97 g of $\text{Li}_{0.93}\text{NbO}_2$ powder and 1.7 g of NOBF_4 in acetonitrile for 10 days. For both reactions, the resultant grey powder was filtered, washed with acetonitrile and ethanol, and dried. Analyses of the X-ray powder diffraction patterns revealed little change from the hexagonal starting material; the pattern obtained for $\text{Li}_{0.50}\text{NbO}_2$ is shown in Fig. 2a. The refined lattice parameters for $\text{Li}_{0.50}\text{NbO}_2$ were $a = 2.919(1) \text{ \AA}$ and $c = 10.453(4) \text{ \AA}$, and those for $\text{Li}_{0.45}\text{NbO}_2$ were $a = 2.9227(9) \text{ \AA}$ and $c = 10.455(5) \text{ \AA}$.

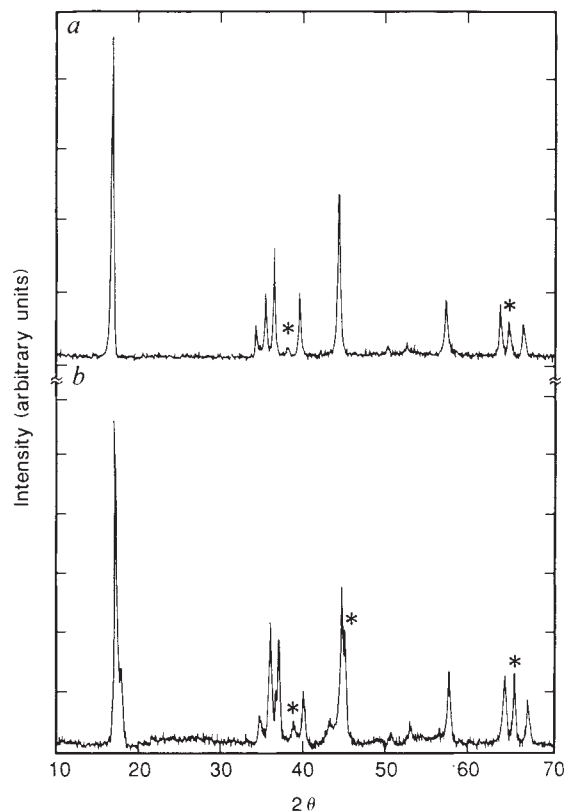


FIG. 2 X-ray powder diffraction patterns of $\text{Li}_{0.50}\text{NbO}_2$ a, before and b, after annealing at 300 °C. Peaks due to the aluminium sample holder are designated with an asterisk.

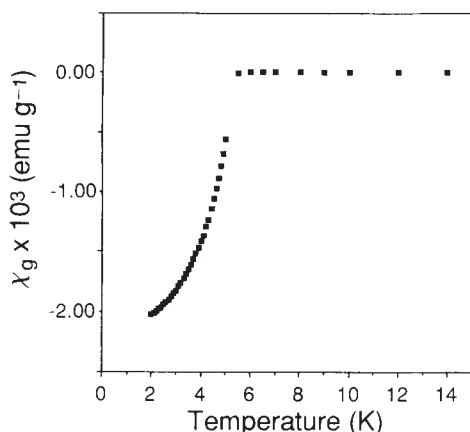


FIG. 3 Meissner effect of $\text{Li}_{0.45}\text{NbO}_2$ prepared by deintercalation of LiNbO_2 with NOBF_4 . Field = 20.33 G.

The magnetic properties of $\text{Li}_{0.50}\text{NbO}_2$ and $\text{Li}_{0.45}\text{NbO}_2$ were measured with a Quantum Design SQUID (superconducting quantum interference device) by cooling in fields of 12.10 G and 20.33 G, respectively. The results for $\text{Li}_{0.45}\text{NbO}_2$ are shown in Fig. 3. Similar behaviour was observed for $\text{Li}_{0.50}\text{NbO}_2$. Above ~6 K both samples behave as temperature-independent paramagnets with very small moments. For $\text{Li}_{0.45}\text{NbO}_2$, the susceptibility drops steeply to diamagnetic values below 5.5 K; this transition indicates the onset of superconductivity. The gram susceptibility at 2 K is $-2.02 \times 10^{-3} \text{ e.m.u. g}^{-1}$. Assuming a density of 6.0 g cm^{-3} and ignoring demagnetization corrections, this corresponds to a Meissner effect of 15%. For $\text{Li}_{0.50}\text{NbO}_2$, T_c onset occurs at the slightly lower temperature of 5 K and the Meissner effect is considerably less. The starting material, $\text{Li}_{0.93}\text{NbO}_2$, and a sample of $\text{Li}_{0.62}\text{NbO}_2$ prepared with less Br_2 so as to yield a larger value of x did not show diamagnetism down to 2 K. These results show that trace amounts of Nb metal, which has a T_c of 9 K, are not present as an impurity in the samples. The large Meissner effect for $\text{Li}_{0.45}\text{NbO}_2$ and the variation of T_c with lithium content provide strong evidence that, when x decreases below a critical value, Li_xNbO_2 becomes a bulk superconductor.

It is useful to compare these results for the oxide with those obtained for Li_xNbS_2 , which has a similar structure (ref. 10). For values of x between 0 and 0.4, the superconducting transition temperature for the disulphide varies between 2 and 6.3 K. But above $x \approx 0.40$, superconductivity has not been observed down to 2 K. For a comparable lithium content, we observe that T_c is substantially higher for the dioxides than for the disulphides. We therefore expect T_c will rise above values seen in the disulphides as the lithium content is reduced further. Unfortunately, with the reagents we used, it is not possible to oxidize Li_xNbO_2 further. We are now investigating different chemical and electrochemical means of reducing the lithium content below $x = 0.45$ so that we can prepare the entire series of Li_xNbO_2 materials and study the relationship between composition and superconductivity.

We have been unable to obtain reliable resistivity data on our samples. Although Kumada *et al.* report resistivity data obtained from pressed powders, we do not know whether the semiconducting behaviour that they observed was due to the sample or to grain-boundary resistances. Furthermore, it was also not possible to sinter the samples to get better contact between the grains. When $\text{Li}_{0.50}\text{NbO}_2$ was heated at 300 °C in an evacuated fused-silica tube for three days, a subtle structural transformation occurred and superconductivity was no longer observed. The X-ray powder diffraction patterns before and after annealing are shown in Fig. 2. A splitting of the most intense peak, indexed as the 002 line in the hexagonal unit cell, indicates lowering of symmetry. Also, the relative intensities of the 004, 100 and 101 lines changed. Differential thermal analysis of the reaction confirmed a structural transformation. There is a strong exotherm beginning about 130 °C and ending at 280 °C indicative of a phase transition; decomposition of this phase occurs near 480 °C. A similar effect was observed on annealing of $\text{Li}_{0.50}\text{VO}_2$, which converts to the LiV_2O_4 spinel phase at 300 °C (ref. 11). We are investigating the exact nature of the structural transformation and the reason for the loss of superconductivity.

The discovery of superconductivity in these materials is significant for a number of reasons. Previously all of the known superconducting niobium oxides contained niobium that was in a low formal oxidation state. In $\text{Li}_{0.45}\text{NbO}_2$ the formal oxidation state of niobium is +3.55. Furthermore, we have seen superconductivity in a layered oxide containing an early transition metal; as the superconducting copper oxides also have layered structures, a comparison of the properties of niobium oxide superconductors with those of the high- T_c copper oxide materials will help clarify the mechanism of superconductivity in layered oxides. □

Received 28 February; accepted 18 April 1990.

1. Bednorz, J. G. & Müller, K. A. *Z. Phys.* **B64**, 189–193 (1986).
2. Wu, M. K. *et al.* *Phys. Rev. Lett.* **58**, 908–910 (1987).
3. Sheng, Z. Z. & Herman, A. M. *Nature* **332**, 55–58 (1988).
4. Parkin, S. S. P. *et al.* *Phys. Rev. Lett.* **60**, 2539–2542 (1988).
5. Johnston, D. C. *et al.* *Mater. Res. Bull.* **8**, 777–784 (1973).
6. Meyer, G. & Hoppe, R. *J. less-common Metals* **46**, 55–65 (1976).
7. Kumada, N. *et al.* *J. Solid St. Chem.* **73**, 33–39 (1988).
8. Burdett, J. K. & Hughbanks, T. *Inorg. Chem.* **24**, 1741–1750 (1985).
9. Bardeen, J., Cooper, L. N. & Schrieffer, J. R. *Phys. Rev.* **B106**, 162–164 (1957); *Phys. Rev.* **B108**, 1175–1204 (1957).
10. McEwen, C. S. & Sienko, M. J. *Rev. chim. Miner.* **19**, 309–316 (1982).
11. de Picciotto, L. A. & Thackeray, M. M. *Mater. Res. Bull.* **20**, 187–195 (1985).

ACKNOWLEDGEMENTS. This work was supported by the U.S. Department of Energy and by the Electric Power Research Institute. We thank C. J. Williams and S. Kos at Perkin Elmer Corporation for the differential thermal analysis, M. Thompson of Raychem Corporation for discussions, and A. King for experimental assistance.

Manufacture of bulk superconducting $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ by a continuous process

R. L. Meng, C. Kinalidis, Y. Y. Sun, L. Gao, Y. K. Tao, P. H. Hor & C. W. Chu

Texas Center for Superconductivity and Department of Physics, University of Houston, Houston, Texas 77204–5506, USA

THE high critical current densities (J_c s) measured in thin films and within grains of the high-transition-temperature (high- T_c) oxide superconductors indicate that low J_c should not prevent the use of high- T_c materials in high-current devices¹. Indeed, a J_c of $\sim 7.5 \times 10^4 \text{ A cm}^{-2}$ has been obtained in highly textured bulk $\text{YBa}_2\text{Cu}_3\text{O}_7$ ('123') at 0 T, processed by a melt-textured² or liquid-phase³ growth technique. Unfortunately, textured materials with this improved J_c comprised only a very small portion (<1 cm) of the bulk precursor, which had random grain orientations on a larger scale. For high-current applications, one needs bulk materials fabricated in usable shapes of large dimensions with controlled grain orientation to carry a high total current, not just a high current density. Here we report the development of a continuous process for fabricating large bulk superconductors with a pre-determined grain orientation. A bar of the 123 compound with dimensions $5 \times 0.5 \times \sim 0.3 \text{ cm}$ with excellent grain alignment (a - b plane along the longest sample dimension) has been fabricated

continuously. The bulk 123 thus obtained has magnetically determined J_c s of $\sim 3 \times 10^4$ and $1.2 \times 10^4 \text{ A cm}^{-2}$ at 0 and 1 T, respectively and transport J_c s of 2×10^4 , 1.1×10^4 and $7.5 \times 10^3 \text{ A cm}^{-2}$ at 0, 0.54 and 0.83 T.

The limited J_c in bulk high- T_c superconductors has been attributed^{4–9} largely to 'weak links' arising from the presence of one or more of the following: (1) high-angle grain boundaries; (2) secondary phases such as Y_2BaCuO_5 , BaCuO_2 and CuO ; (3) porosity; (4) local or global deficiency of oxygen; (5) microcracks; (6) the electrical anisotropy. To overcome some of these problems, various melt-texturing techniques have been used.

As a 123 bulk precursor cools from a single-phase liquid region to $\sim 1,050^\circ\text{C}$ it undergoes the transition ($\text{Y}_2\text{O}_3 + \text{liquid}$) $\rightarrow$ $\text{Y}_2\text{BaCuO}_5 + \text{liquid}$ $\rightarrow$ ($123 + \text{liquid of composition BaCuO}_2 + \text{CuO}$) (refs 10, 11). Near $\sim 1,050^\circ\text{C}$, 123 forms through a peritectic solidification according to $\text{Y}_2\text{BaCuO}_5 + \text{liquid}$ ($3\text{BaCuO}_2 + 2\text{CuO}$) $\rightarrow 2\text{YBa}_2\text{Cu}_3\text{O}_{6.5}$. Heating the material above the solidus temperature ($1,020$ – $1,200^\circ\text{C}$) helps to dissolve the impurity phases, which are precipitated in the grain boundaries. Also, crystal grains grow more easily in a liquid solution. Therefore, the reported methods^{2,3,12,13} for making 123 that is capable of carrying a high J_c consist of rapidly heating 123 above the solidus temperature (or even above the melting point) and slowly cooling through the peritectic temperature. This has been achieved using a prescribed temperature schedule, that is, different specified rates of temperature change (dT/dt) for various temperature regions. The morphology of the grains obtained depends critically on the temperature schedule^{2,14}, although in a complex manner that remains to be clarified. Some workers² have obtained ordered grain growth by directional solidification using a temperature gradient along the length of the material, dT/dx , whereas others³ have approached the problem without using temperature gradients. After completion of our work, it came to our attention that McGinn *et al.*^{15,16} have carried out a texture processing of bulk 123 by zone melting. They did not use a prescribed temperature gradient and achieved only small domains of aligned grains, which had a magnetically determined J_c of $\sim 1.8 \times 10^4 \text{ A cm}^{-2}$ at 2 T and a transport J_c of $\sim 500 \text{ A cm}^{-2}$ at 0.5 T. To allow the continuous growth of large samples of bulk 123 that are capable of carrying a high J_c and to help remove impurities, we move the 123 bulk precursor through an open tubular furnace having a pre-set temperature profile that includes a narrow hot zone (Fig. 1). We vary both the speed, dx/dt , and the gradient, dT/dx , to give the desired rate of temperature change, dT/dt .

A bar of dimensions $5.2 \times 0.5 \times 0.3 \text{ cm}$ was sintered from a 123 fine-powder precursor at 940 – 960°C for 24 h in an oxygen atmosphere to obtain $\geq 90\%$ of the theoretical density, 6.3 g cm^{-3} . The bar was heated to $\sim 1,100^\circ\text{C}$ very rapidly for a short time before it was cooled rapidly to near the peritectic temperature ($\sim 1,050^\circ\text{C}$). The sample was then allowed to undergo peritectic solidification very slowly. Because of the narrow hot zone, zone refining occurs along the axis of travel depositing Y_2BaCuO_5 at one end of some of the samples. Pieces from different parts of the bar were oxygenated at 500°C for 24–48 h. Figure 2 shows representative scanning electron micrographs of the longitudinal and transverse fracture faces of the pieces. There is excellent grain alignment throughout the bar. The X-ray diffraction patterns in Fig. 2 clearly show that the a - b plane coincides with the longitudinal fracture face along the direction of travel and that all of these planes in the sample are parallel to each other. The electrical and magnetic susceptibility measurements (Fig. 3) exhibit sharp transitions, demonstrating the high superconducting quality of the sample. The small Meissner signal determined during field cooling compared with that obtained during zero-field cooling (Fig. 3) indicates the strong flux pinning existing in those samples. According to the Bean formula ($J_c = 20 \Delta M/d$ where ΔM is the difference in the remnant moments of the sample during the increasing and decreasing field sweeps, and d is the sample dimension

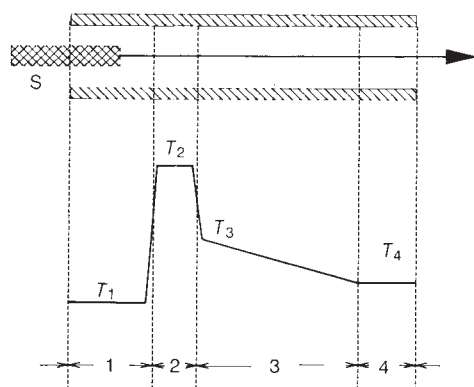


FIG. 1 The schematic of the continuous process for step (3) in the preparation of a bulk 123 sample (S). The temperatures are: $T_1 = 500 \pm 425^\circ\text{C}$, $T_2 = 1,150 \pm 50^\circ\text{C}$, $T_3 = 1,035 \pm 15^\circ\text{C}$ and $T_4 = 925 \pm 25^\circ\text{C}$. The times spent in different zones are: 1, not critical; 2, $10 \pm 5 \text{ min}$; 3, $100 \pm 50 \text{ h}$; 4, $20 \pm 5 \text{ h}$. The peritectic temperature is $\sim 1,050^\circ\text{C}$.

FIG. 2 Representative scanning electron micrographs of the *a*, longitudinal and *b*, transverse fracture surfaces of the bar of 123 compound, together with their X-ray diffraction patterns.

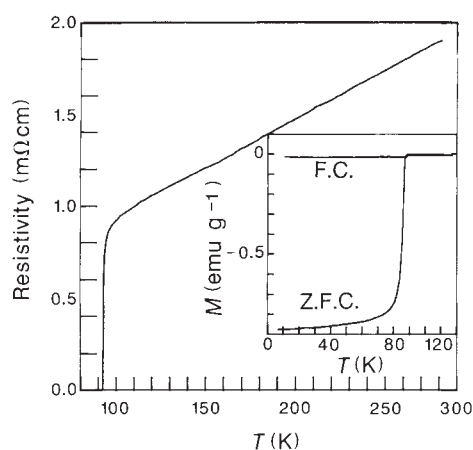
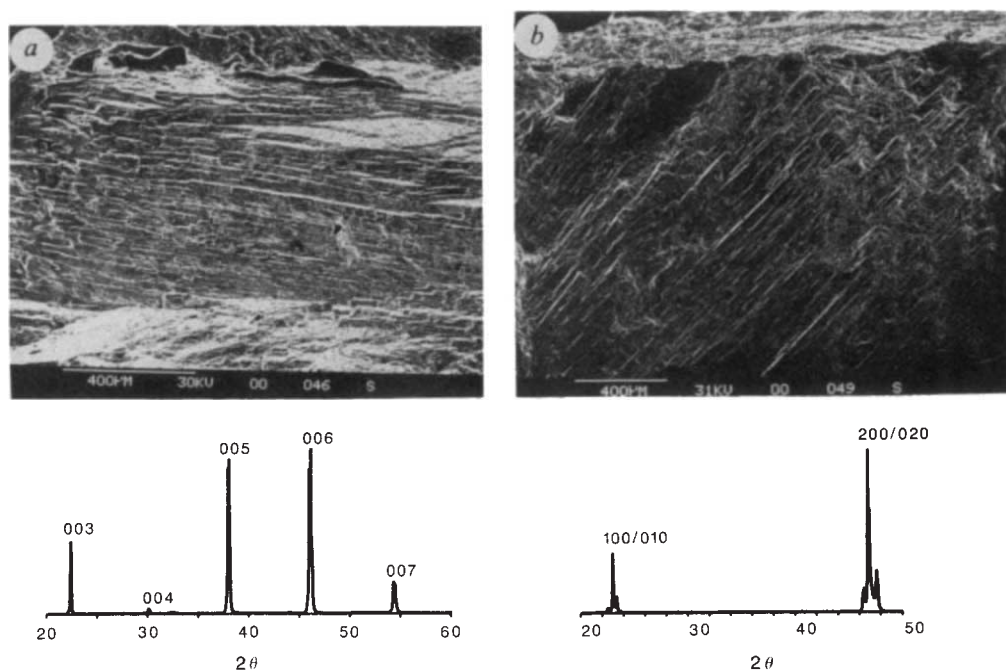


FIG. 3 Resistive and magnetic transitions of the continuously processed 123 (F.C., field cooling; Z.F.C., zero-field cooling).

perpendicular to the field) the J_c of a sample of dimensions $0.216 \times 0.165 \times 0.165$ cm (that is, $d = 0.165$) is 3×10^4 and 1.2×10^4 at 0 and 1 T, respectively. At 77 K and with the magnetic field perpendicular to J_c the transport J_c (Fig. 4) in samples of dimensions $\sim 1 \times 0.04 \times 0.05$ cm was $\sim 2 \times 10^4$ A cm $^{-2}$ at 0 T, 1.1×10^4 A cm $^{-2}$ at 0.54 T and 7.5×10^3 A cm $^{-2}$ at 0.82 T. When the magnetic field is parallel to J_c , $J_c \sim 1.4 \times 10^4$ A cm $^{-2}$ at 0.82 T, suggesting that weak links are not a serious problem in materials processed in this way. Transport J_s as low as 3×10^3 A cm $^{-2}$ at 0 T, however, were detected in one piece of the sample. We attribute the difference to chemical inhomogeneity. For example, the oxygen content (determined by a Raman microprobe) varied from 6.92 to 6.87. We believe that adjusting the processing parameters should give a material with a higher, more uniform transport J_c .

Because of the continuous nature of the process, desired sizes of practical forms (such as plates, rods, wires, ribbons and even thick films) could be made from highly textured bulk superconductors having high J_c s. The processing of the bulk 123 can be divided into four steps: (1) synthesis of powder precursor; (2)

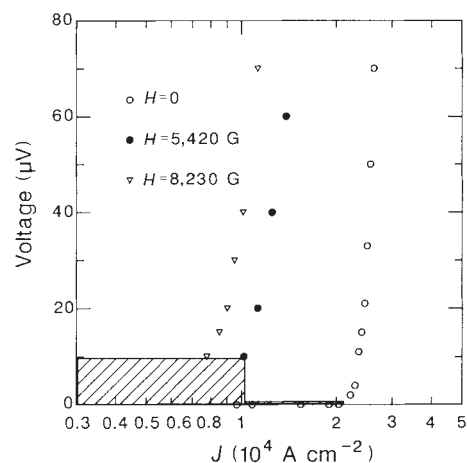


FIG. 4 Voltage as a function of current density at different magnetic fields (perpendicular to J_c) at 77 K. The resolution was 0.2 μ V and 10 μ V, respectively, at a magnetic field, H , of 0 and >0.4 T, in contrast to the 50 μ V previously achieved³.

sintering of bulk precursor; (3) melt-textured growth; (4) oxygenation. By adopting the proper temperature gradients and speeds in different zones of the furnace, steps (2)–(4) can be combined into a single continuous process. The processing temperature, particularly in step (3), can be lowered by using low-melting flux consisting of BaCuO $_2$ and/or CuO, a more reducing atmosphere, or halogen¹⁷. This is especially crucial for wires or ribbons of superconductors having metal cladding such as a silver sheath. At present, the process is slow but by varying the processing parameters and possibly the phase diagram of Y–Ba–Cu–O through doping, we expect to improve the speed of the grain growth. □

Received 6 March; accepted 30 April 1990.

1. Chu, C. W. *Science* **245**, 111–112 (1989).
2. Jin, S. *et al. Appl. Phys. Lett.* **42**, 2074–2076 (1988).
3. Salama, K., Selvamaniyam, V., Gao, L. & Sun, K. *Appl. Phys. Lett.* **54**, 2352–2354 (1989).
4. Dinger, T. R., Worthington, T. K., Gallagher, W. J. & Sandstrom, R. L., *Phys. Rev. Lett.* **58**, 2687–2690 (1987).
5. Nakahara, S., Fisanick, G. J., Yan, M. F., van Dover, R. B. & Boone, T. *J. Cryst. Growth* **85**, 639–650 (1987).

6. McGuire, T. R. *et al.* *Phys. Rev.* **B36**, 4032-4035 (1987).
7. Tozer, S. W., Kleinsasser, A. W., Penney, T., Kaiser, D. & Holtzberg, F. *Phys. Rev. Lett.* **59**, 1768-1771 (1987).
8. O'Bryan, H. M. & Gallagher, P. K. *Adv. Ceram. Mater.* **2**, 640 (1987).
9. Ekin, J. W. *et al.* *J. appl. Phys.* **62**, 4821-2828 (1987).
10. Aselage, T. & Keefer, K. *J. Mater. Res.* **3**, 279-291 (1988).
11. Murakami, M., Morita, M., Doi, K., Miyamoto, K. & Hamada, H. *Jap. J. appl. Phys.* **28**, L399-L401 (1989).
12. Murakami, M., Morita, M., Doi, K. & Miyamoto, K. *Jap. J. appl. Phys.* **28**, 1189-1194 (1989).
13. Alford, N. M. *et al.* *J. appl. Phys.* **66**, 5930-5934 (1989).
14. McGinn, P. J., Black, M. A. & Valenznek, A. *Physica C156*, 57-61 (1988).
15. McGinn, P. J., Chen, W. & Black, M. A. *Physica C161*, 198-203 (1989).
16. McGinn, P. J., Chen, W., Zhu, N., Balachandran, U. & Lanagan, M. T. *Physica C165*, 480-484 (1989).
17. Radousky *et al.* Preprint (Lawrence Livermore National Laboratory).

ACKNOWLEDGEMENTS. We thank J. H. Choi and D. Pham for technical help. This work is supported in part by DARPA, NASA, the Texas Center for Superconductivity at the University of Houston, the NSF and the T. L. L. Temple Foundation.

Surface-adsorbed free radicals observed by positive-muon avoided-level-crossing resonance

Ivan D. Reid*, Toshiyuki Azuma* & Emil Roduner

Physikalisch-Chemisches Institut der Universität Zürich,
Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

HETEROGENEOUSLY catalysed reactions usually proceed at outer or inner surfaces of materials that have large and often non-uniform surface area. Many intermediates of such chemical transformations are short-lived and escape direct detection; such transient species may include free radicals. The structure, adsorption and surface dynamics of diamagnetic intermediates have been well studied by thermodynamic or spectroscopic methods, but paramagnetic species (that is, radicals) recombine at catalytically relevant temperatures so that they are never present in sufficiently high concentrations to allow their observation by these techniques. The otherwise powerful laser methods are also hampered by the non-transparent nature of heterogeneous systems. Here we demonstrate for the model case of cyclohexadienyl radicals on silica that a novel technique, positive-muon avoided-level-crossing spin resonance, which uses positive muons as spin probes, is sufficiently sensitive to allow the detection and study of radicals under such constrained conditions. The technique can detect radical concentrations many orders of magnitude less than can be studied by conventional magnetic spectroscopy.

Spin-polarized energetic positive muons injected into matter are used routinely to sense magnetic interaction phenomena in various environments¹. As they thermalize in the sample a certain fraction form muonium atoms ($\text{Mu} \equiv \mu^+e^-$) by electron capture near thermal energy. Mu behaves as a very light isotope of hydrogen ($m_{\text{mu}} \approx \frac{1}{1836}m_{\text{H}}$) and undergoes similar chemical reactions. In particular, it adds to unsaturated bonds to form muonated free radicals in which the muon is a spin-polarized substitute for a normal hydrogen nucleus²; for example, addition to benzene leads to the muonated cyclohexadienyl radical $\text{C}_6\text{H}_6\text{Mu}$. A single-particle-counting technique that takes advantage of the asymmetry of the muon decay ($\tau_{\mu} = 2.2 \mu\text{s}$, $\mu^+ \rightarrow e^+ + \bar{\nu}_{\mu} + \nu_e$, the positron being emitted preferentially along the instantaneous muon spin axis) reveals the evolution of spin polarization. In the conventional variant of the experimental technique, with a magnetic field transverse to the initial spin polarization¹, temporal variations in the arrival rate of decay positrons in a given direction show the precession frequencies of the muons; the signal is analogous to a free-induction decay in pulsed magnetic resonance. This method has been used to study the radiolytic formation process and the structure of

radicals in liquids, and to determine accurate absolute rate constants of fast radical reactions which are not easily accessible by conventional means².

We report experiments using a type of muon spectroscopy based on the effect of avoided level crossings^{3,4}. Figure 1 illustrates magnetic energy levels of a three-spin system (μ^+ , e^- , and nucleus k , chosen here to be a proton) at high magnetic fields B . Over most of the field range the energy eigenstates are stationary Zeeman states (denoted by α or β according to the orientation of the spins to the field). Crossing of states may be avoided by Fermi contact or dipolar interaction of the spins. In these regions, eigenstates are oscillating mixtures of pure Zeeman states. Therefore, an oscillatory inversion of the muon spin along the field results whenever the interaction involves one state with muon spin α_{μ} and one with β_{μ} . The inversion is seen experimentally by monitoring changes in the muon spin orientation when a magnetic field longitudinal to the spin polarization of the muon beam impinging on the sample is scanned through the region of an avoided crossing. Resonances are observed as changes in decay positron count rates in detectors placed in the 'forward' and 'backward' directions relative to the initial spin polarization. The muon polarization is then proportional to the experimental asymmetry $\mathcal{A}$, where

$$\mathcal{A} = \frac{\sum N_f - \sum N_b}{\sum N_f + \sum N_b} \quad (1)$$

and $\sum N_f$ and $\sum N_b$ are the total number of positrons detected in the forward and backward directions respectively. Signals appear as dips in $\mathcal{A}$ as a function of field because the resonant process transfers muon polarization from the forward to the backward direction, enhancing $\sum N_b$ at the expense of $\sum N_f$.

Prominent resonances correspond to the magnetic selection rule $\Delta M = 0$, where $M = m_e + m_{\mu} + m_k$ is the total spin projection, and represent the exchange of spin polarization between the muon and a magnetic nucleus. To first order they occur at fields

$$B = \left| \frac{A_{\mu} - A_k}{2(\gamma_{\mu} - \gamma_k)} \right| \quad (2)$$

where A_i are the hyperfine couplings and γ_i the gyromagnetic ratios. Here, crossing of the levels is avoided only indirectly by coupling of both states to an energetically distant third state, but the isotropic Fermi contact interaction suffices to lift the degeneracy of energies⁵. A fundamentally different type of resonance with $\Delta M = 1$ occurs around

$$B = \left| \frac{A_{\mu}}{2\gamma_{\mu}} \right| \quad (3)$$

In contrast to the $\Delta M = 0$ resonance, the crossing of levels is

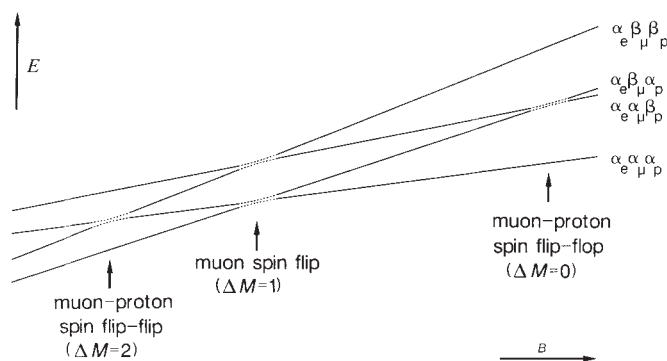


FIG. 1 High-field energy-level diagram for a three-spin system of an electron e , positive muon μ , and proton p . States are labelled by the spin projections α or β along the field B . Avoided-level-crossing resonances occur when states with opposite muon spins become near-degenerate in energy, allowing the system to oscillate between them.

* Present addresses: Paul Scherrer Institute, CH-5232 Villigen PSI, Switzerland (I.D.R.); Institute of Physics, College of Arts and Sciences, University of Tokyo, Komaba 3-8-1, Meguro, Tokyo, 153 Japan (T.A.).

avoided directly by coupling of two Zeeman states through the dipolar part of the hyperfine interaction⁴. Phenomenologically, this resonance can be viewed as muon precession in the remnant transverse component of the local field when the external field cancels the longitudinal component, so under isotropic conditions the resonance is lost. This is the case in a liquid, for example, where a radical averages dynamically over all orientations with the critical time being set by the hyperfine anisotropy, typically 10–30 MHz. Where mobility is restricted, as in a solid, this transition is expected to be observable and we have found in bulk benzene that it sets in extremely strongly below the freezing point⁶. A third type of resonance with $\Delta M = 2$ only appears under special orientations of the hyperfine coupling tensors to the magnetic field, and will not be considered further here.

The muonated cyclohexadienyl radical adsorbed on silica has been observed using the transverse-field muon-spin-rotation technique⁷. However, as radical formation was slow there was a dephasing of the muons forming radicals owing to significant precession in the precursor state. Thus the signal amplitude was greatly reduced and it was not feasible to determine reliable line widths as a function of temperature. The avoided-level-crossing technique is not as sensitive to formation rate because there is no precession in longitudinal fields, so radicals may be detected under much more dilute conditions with correspondingly longer radical formation times. Furthermore, hyperfine couplings are usually obtained for all magnetic nuclei in the radical, whereas the transverse-field technique provides only the muon coupling and thus a less detailed characterization of the species.

Our experiments were done on a SiO₂ powder sample of Cab-O-Sil grade EH-5 with an average grain diameter of 7 nm and a surface area of 380 m² g⁻¹ (Cabot Corp., technical report). The powder (6 g) was placed in a stainless-steel cell with a muon entrance window of 50- μ m stainless foil. The cell was baked under vacuum for 48 h at $\sim$ 390 K to remove adsorbed water, leaving $\sim$ 4 hydroxyl groups nm⁻² at the surface (Cabot Corp., technical report). Based on a molecular cross-section of 0.40 nm², 0.75 g of degassed benzene was condensed into the cell to form a monolayer coverage⁸, and the vacuum connection then sealed. Measurements were made using muons produced at the accelerator of the Paul Scherrer Institute. The sample was placed in the centre of a magnet of variable longitudinal field

and held at a controlled temperature during measurements. Counts of decay positrons were accumulated at each field value for a period of $\sim$ 30 s and the asymmetry $\mathcal{A}$ calculated as above.

At 334 K we observed a strong avoided-level-crossing resonance at 2.070 T and two weaker ones at 2.885 T and 2.945 T. They are all of the $\Delta M = 0$ type, involving different protons, and identify the muonated cyclohexadienyl radical which in the liquid has $A_{\mu} = 513$ MHz and $A_k = 126$ MHz (methylene proton), -25 MHz (ortho protons), $+7.5$ MHz (meta protons) and -36 MHz (para proton)⁹. Equation (2) thus associates the resonances with the methylene, ortho and para protons respectively. The methylene proton signal was studied at 12 temperatures in the range 139–334 K. Figure 2 shows representative signals where a slowly varying background due to changes in beam optics with magnetic field has been subtracted from the data. The change in resonance position is a reflection of the temperature dependence of the hyperfine couplings. The line shape is lorentzian at all temperatures, signifying isotropic conditions. We note that the $\Delta M = 1$ resonance around 1.9 T, which is a good indicator of anisotropy, was not detected at any temperature, so we can immediately conclude that there is no residual anisotropy in the C₆H₆Mu/Cab-O-Sil system over the temperature range studied.

Comparison of the signal amplitude with that observed in bulk liquid benzene reveals that about two-thirds of the muons reach a C₆H₆Mu radical state in the Cab-O-Sil sample at 334 K. The yield decreases continuously with temperature until the signal disappears below 139 K. We ascribe this to the slower thermal diffusion of the muons out of the silica grains and a lower radical formation rate at the lower temperatures.

The heat of adsorption of benzene onto a fully hydroxylated silica surface is 45 kJ mol⁻¹, attributed to polarization of the benzene π -orbitals by the hydroxyl groups¹⁰. In the absence of any evidence to the contrary, it is reasonable to assume that the same binding mechanism is true for the radical. The avoided-level-crossing resonance is shifted to higher fields, relative to the liquid, by about 5 mT at comparable temperatures, corresponding to an increase in the hyperfine coupling constants of the two methylene nuclei (muon and proton) of $\sim$ 0.4%. An increase in the muon coupling was also seen in the transverse-field experiment⁷, so the environment of the radical is evidently different from that in the liquid. A similar effect has been observed by electron spin resonance for the methyl radical adsorbed on vycor in comparison to *n*-hexane solution¹¹.

The line width is 7.6 mT HWHM (half width at half maximum) at the highest temperature. This exceeds the intrinsic width of 6.04 mT by only 25% and is close to the value observed in liquid benzene, illustrating similar mobilities for the radicals in the two environments at this temperature. The width increases continuously with decreasing temperature, in total by a factor of five over the whole range, indicating that the fluctuation of the hyperfine interaction becomes slower due to decreasing mobility of the radical. Three types of motion could be responsible for line-width effects. The first is rotation about an axis perpendicular to the surface, leading to a cylindrical Hamiltonian by partial averaging of the hyperfine anisotropy. It has been demonstrated for such a case in a plastic crystal that the $\Delta M = 1$ resonance is very strong even for small anisotropies¹². The absence of this resonance at all temperatures means that the motion is not simply uniaxial. The second type of motion is rotation about an axis parallel to the surface. For a planar species like C₆H₆Mu this is more likely a flip, that is a jump between two orientations with identical hyperfine interactions, which also would not lead to the disappearance of the $\Delta M = 1$ resonance. The third motion to be considered is translation. On a flat surface this would not affect the orientation of the radical with respect to the magnetic field, but for a more or less spherical grain with a non-uniform surface it would very quickly produce an average over all orientations and hence the loss of the $\Delta M = 1$ signal.

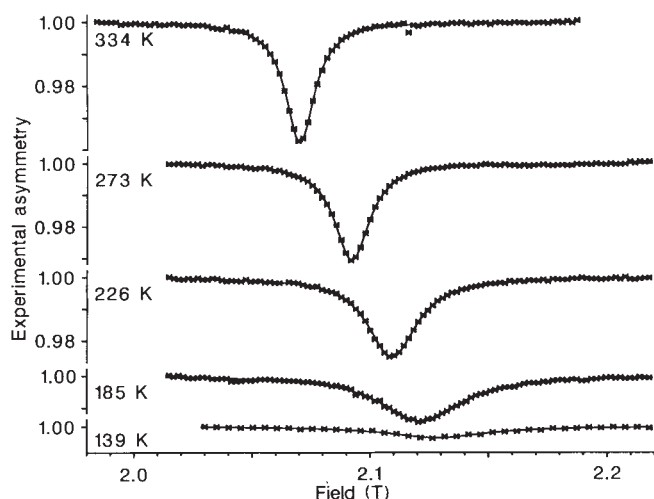


FIG. 2 $\Delta M = 0$ avoided-level-crossing resonances obtained from the methylene proton in muonated cyclohexadienyl radicals in a monolayer of benzene on 7-nm SiO₂ grains. The data were fitted to a Lorentzian resonance on a cubic background, which has been removed in these displays; the solid curves give the fitted resonances.

Translation seems to be the only type of motion that can explain our results. Clearly, the $\Delta M = 0$ resonance shows line-width effects because it has a different dependence on hyperfine couplings and senses motion on a different timescale to the $\Delta M = 1$ transition. Models for quantitative analysis of line broadening exist so far only for chemical reactions and spin exchange effects¹³ but will have to be developed for reorientational motion. Qualitatively, our present observations show that despite a reduction in mobility at low temperatures, complete randomization of orientation of cyclohexadienyl radicals on a silica surface is still achieved in ≤ 10 ns. Furthermore, it is interesting to note that over the temperature range studied there is no evidence of a phase transition in the adsorbed benzene film such as, for example, two-dimensional melting which has been observed for a benzene monolayer on graphite at ~ 130 K (ref. 14).

The main advantage of muon spectroscopy for studying surface-adsorbed radicals lies in the sensitivity of the technique. The high degree of spin polarization ($\sim 100\%$) and the single-particle detection mode allow it to work with concentrations down to just one muonated radical at a time in the entire sample. Electron-spin-resonance detection, for comparison, needs $\sim 10^{12}$ radicals in the cavity, corresponding to a concentration which usually cannot be maintained at temperatures where the radicals become mobile, as high mobility allows the radicals to diffuse together and terminate by combination (for example, to dicyclohexadiene in this case) or disproportionation (for example, to

benzene and cyclohexadiene). The non-transparent and inhomogeneous nature of most heterogeneous catalysts on the other hand does not permit the use of optical methods. The avoided-level-crossing technique presented here seems to be suitable to fill the gap, and it is expected that it will allow the investigation of structure and dynamic behaviour of radicals on further materials of catalytic interest, such as zeolites and metal-loaded oxides, and that theory will be developed to permit the extraction of correlation times and activation energies from experimental data. $\square$

Received 22 January; accepted 14 March 1990.

1. Schenck, A. *Muon Spin Rotation Spectroscopy. Principles and Applications in Solid State Physics*. (Adam Hilger, Bristol, 1985).
2. Roduner, E. *The Positive Muon as a Probe in Free Radical Chemistry* Lecture Notes in Chemistry 49 (Springer, Berlin, 1988).
3. Abragam, A. *C. r. hebdom. Séanc. Acad. Sci., Paris II* **299**, 95–99 (1984).
4. Kieft, R. F. *Hyperfine Interact.* **32**, 707–720 (1986).
5. Heming, M. *et al. Chem. Phys. Lett.* **128**, 100–106 (1986).
6. Roduner, E. *Chimia* **43**, 86–97 (1989).
7. Heming, M. Z. *Phys. Chem. Neue Folge* **151**, 35–50 (1987).
8. Gregg, S. J. & Sing, K. S. W. *Adsorption, Surface Area and Porosity* (Academic, London, 1967).
9. Percival, P. W. *et al. Chem. Phys. Lett.* **133**, 465–470 (1987).
10. Kiselev, A. V. & Lygin, V. I. *Infrared Spectra of Surface Compounds* (Wiley, New York, 1975).
11. Fischer, H. & Hefter, H. Z. *Naturf.* **23a**, 1763–1765 (1968).
12. Roduner, E., Reid, I. D., Riccio, M. & De Renzi, R. *Ber. Bunsenges. phys. Chem.* **93**, 1194–1197 (1989).
13. Heming, M. *et al. Chem. Phys.* **129**, 335–350 (1989).
14. Boddenberg, B. & Grosse, R. Z. *Naturf.* **42a**, 272–274 (1987).

ACKNOWLEDGEMENTS. Support from the Swiss National Foundation for Scientific Research and by the Paul Scherrer Institute is gratefully acknowledged. We thank Theo Hohl for assistance with the experiments and Professor H. Fischer for his continuing interest and support.

Efficient photoelectrochemical solar cells from electrolyte modification

Stuart Licht & Dharmasena Peramunage

Department of Chemistry, Clark University, Worcester, Massachusetts 01610, USA

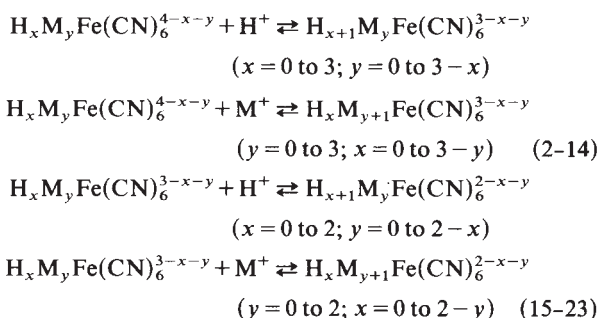
PHOTOELECTROCHEMICAL solar cells (PECs)^{1–3} have shown energy conversion efficiencies approaching 13% in sunlight⁴, and up to 15% in simulated insolation⁵. Of these, only those incorporating n-cadmium chalcogenide electrodes have been demonstrated to be conducive to thin film⁶ or *in situ* storage systems⁷. Previous studies of photoelectrochemical current and voltage limitation^{2,3,5,8,9} have focused on modification of the semiconductor electrode. Here we take the alternative approach by demonstrating that energy conversion can be improved by prevention of electrode surface modification and by systematic modification of the electrolyte. Electrolyte modification entails investigations of the primary photo-oxidized species, the nature of the counter ion, the distribution of species in solution, and related competing reactions. Optimization of the distribution of species and addition of cyanide to n-CdSe/([KFe(CN)₆]^{2–/3–})_{aq} PECs enhances the available voltage and the ease of charge transfer, and suppresses related decomposition products. The resultant PEC achieves an open-circuit potential of 1.2 V, an efficiency of 16.4%—the highest for any wide-band-gap (1.7 eV) solar cell (solid state or photoelectrochemical)—and a 100-fold improvement in photocurrent lifetime. Each of these represents a step towards realization of a viable PEC.

The highest conversion efficiencies in n-CdSe/[Fe(CN)₆]^{3–/4–}_{aq} solar cells, 12 to 14%, have been reported using electrolytes containing a 1:1 ratio of Fe(CN)₆^{4–} to Fe(CN)₆^{3–} salts in a highly alkaline environment^{8,10}. The observed photocurrents are, however, highly unstable and decay within hours⁶. Attempts to enhance the photoconversion stability of these systems have included cationic surface modification of the n-CdSe photoelectrode⁹ and isolation of the wavelength dependence of the surface instability¹¹.

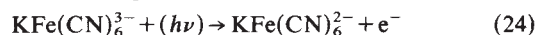
Photo-oxidative processes in n-CdSe/Fe(CN)₆^{3–/4–} PECs are typically represented as



The role of the various redox-active constituents in constraining n-CdSe/Fe(CN)₆^{3–/4–} photo-oxidative energy conversion has not been previously investigated, however, and ferricyanide and ferrocyanide salts show complex aqueous equilibria. In a series of studies on n-cadmium chalcogenide aqueous polysulphide PECs, we have emphasized the importance of isolating the active electrolytic constituents and identifying the redox reactions pertinent to overall energy conversion¹². This approach is expanded in this study of n-CdSe ferri/ferrocyanide PECs. Monovalent cation salts of ferri and ferrocyanide, M₃Fe(CN)₆ and M₄Fe(CN)₆ can include the 22 aqueous equilibria relations:



Thermodynamic constraints for some of these equilibria have been investigated^{13–15}. Our numerical iterative analysis, applying these known equilibria for M = K in the high-pH electrolytes we use here, shows that the dominant species in solution are the monocationic species KFe(CN)₆^{3–} and KFe(CN)₆^{2–}, and the photoinduced oxidative process may be represented



Equation (24) is consistent with the findings¹⁶ that electrochemical ferrocyanide oxidation (in the dark) is first order with respect to cation, and that potassium and caesium salts are electrocatalytically superior to the other alkali ferrocyanides. Here we have used the potassium salts.

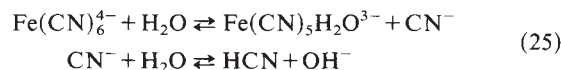
An improvement in measured photocurrent stability with

increase in pH is seen in comparing pH 3.8 and pH 13.3 electrolytes (Fig. 1, curves a, b). Photocurrent stability of about one hour is observed in the high-pH electrolyte (Fig. 1, curve b), comparable to the best previous n-CdSe/Fe(CN)₆^{3-/4-} PEC stability⁹. Our measurements at high pH and at constant illumination, for a wide variation of K₄Fe(CN)₆ and K₃Fe(CN)₆ concentrations in n-CdSe/Fe(CN)₆^{3-/4-} resulted in PECs that yield similar (to within ±2%) short-circuit photocurrents, but widely varying stability. Methods to measure light-pulsed transient photocurrents have been developed to probe PEC kinetics¹⁷ but the slow (ms) response of that methodology provides only a qualitative measure of interfacial kinetics¹⁸. Here we have measured, under constant photon flux, potential-stepped transient photocurrent (PSTP) time response as a probe of PEC kinetics. PSTP is comparable to or faster than the interfacial double-layer charging time, and is only limited by electronic time constraints (20 μs) of the potentiostat. We compared PSTP time response of an n-CdSe electrode immersed in aqueous electrolytes containing constant concentration of K₄Fe(CN)₆ as the K₃Fe(CN)₆ concentration was varied by three orders of magnitude. It is evident that charge transfer is most facile at a ratio of ~20:1 between KFe(CN)₆³⁻ to KFe(CN)₆²⁻ salts (Fig. 2). We found that this optimal ratio is not sensitive to changes in pH, total salt or K₄Fe(CN)₆ concentrations.

The transient photocurrent measurements indicate that at a KFe(CN)₆³⁻:KFe(CN)₆²⁻ concentration ratio of ~20:1, more photogenerated carriers should participate in the constructive solution redox process (equation (24)) with fewer photogenerated carriers participating in photocorrosive processes. We do observe a substantial enhancement in photocurrent stability for the n-CdSe/KFe(CN)₆^{2-/3-} system for this concentration ratio (Fig. 1, curves b, c); other electrolytes resulted in decreased photocurrent stability.

In addition to equations (2–23), a variety of other redox and substitution reactions contribute to the aqueous ferri/ferrocyanide system. In saturated phase and low-pH systems, the relatively insoluble cyanic acids and Prussian blue salts dominate¹⁹. Increase of solution pH substantially enhances n-CdSe/KFe(CN)₆^{2-/3-} PEC conversion efficiencies and cell stability. Various related properties were investigated, but the basis of this pH effect was not understood¹⁰. We propose that this suppression of photocorrosion and increase in stability are secondary effects that are the result of photoinduced electrolyte (as opposed to electrode) instability. Electrolyte photodecompo-

sition occurs according to²⁰

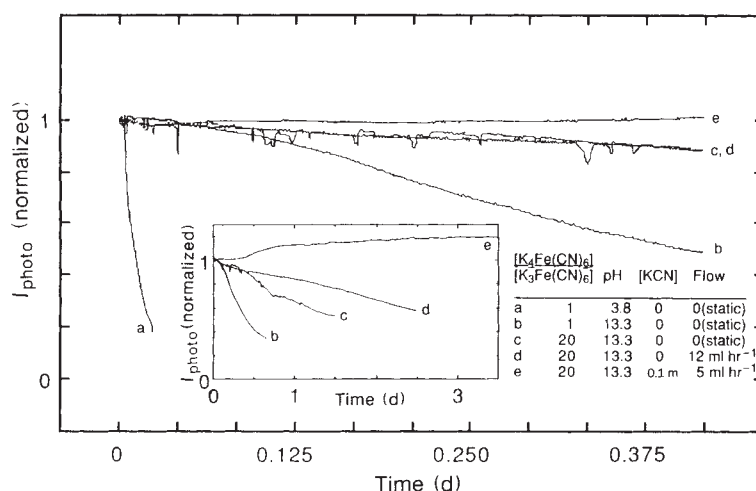


The energy of activation of electrolyte photodecomposition is 19.8 kcal mol⁻¹; the rate is maximum at pH ~4 and falls off rapidly with increasing pH²⁰. Fe(CN)₅H₂O³⁻ has been shown to be highly reactive²¹. On illumination of the n-CdSe/KFe(CN)₆^{2-/3-} PECs, we detect spectroscopically significant consumption of KFe(CN)₆³⁻ and significant cyanide generation in solutions at pH 3.8 (adjusted by HCl addition), and none in solutions at pH > 13 (adjusted by addition of KOH). This is consistent with the poor stability and rapid photoinduced solution decomposition in the low-pH electrolyte (Fig. 1, curves a, b). When we use the pH 3.8 solution as a photoelectrolyte for the PEC, a brown photoelectrode overlayer forms within seconds, consistent with known KCdFe(CN)₆ and Cd₄[KFe(CN)₆]₃ salts^{22,23}. The observed rate of overlayer formation diminishes by several orders of magnitude in the high-pH 0.25 m K₄Fe(CN)₆/0.0125 m K₃Fe(CN)₆ electrolyte. Consumption of Fe(CN)₆⁴⁻, equation (25), will diminish availability of the (photo)active hexacyanide species leading to reduced PEC performance, equation (24). The effect of photoinduced decomposition of electrolyte on the long-term stability of n-CdSe/Fe(CN)₆^{2-/3-} photoelectrochemical cells is readily detected (Fig. 1, curves c, d). Further improvement in photocurrent stability occurs when electrolyte is renewed by continuously removing photoinduced electrolyte decomposition products from the cell (Fig. 1 inset, curve d).

Consistent with the stability measurements (Fig. 1, curves c, d), added cyanide may be expected to slow down the electrolyte decomposition by enhancing the reverse reaction, equation (25). The presence of CN_{aq}⁻ also affects solubility. Addition of Cd²⁺_{aq} (as dissolved CdCl₂) to alkaline K₃Fe(CN)₆/K₄Fe(CN)₆ electrolytes immediately precipitates insoluble cadmium ferri/ferrocyanide salts. But when 0.1 m KCN is incorporated in the same electrolytes added Cd²⁺_{aq} is entirely soluble up to a CdCl₂:KCN concentration ratio of 1:4, dissolving as the complex Cd(CN)₄²⁻_{aq}. Crystalline CdSe remains insoluble in either electrolyte. Cyanide can therefore remove otherwise insoluble alkali cadmium ferri/ferrocyanide photodecomposition products⁹, preventing surface passivation and maintaining an active clean CdSe surface morphology.

Further dramatic improvements in photocurrent stability are

FIG. 1 Photocurrent stability of single-crystal n-CdSe immersed in a variety of 0.25 m K₄Fe(CN)₆ solutions. Inset: Long-term stability. 0.200-cm², 1-ohm-cm n-CdSe (Cleveland Crystal) was mounted with the (1120) plane exposed, and was pretreated with successive 5-μm, 1-μm, 0.35-μm and 0.05-μm mechanical polishing, chemically etched in 2% bromine/methanol (by volume) and photoetched⁴ in 1.0 m Na₂SO₄. Electrolytes were freshly prepared with analytical grade reagents (Aldrich Chemicals) in deionized distilled water. Before use all electrolytes were purged with high-purity argon. Intensity of the 100-W tungsten-halogen illumination was initially varied by distance to provide an illumination of ~25 mW cm⁻² for data presented in curves a to d and ~40 mW cm⁻² for data presented in curve e, calibrated by outdoor comparison to air mass 1.5 insolation. Photocurrents measured in curves a–e were normalized relative to the initial photocurrents (initial photocurrents of 2.10, 2.97, 2.79, 3.12 and 6.77 mA cm⁻² were measured for data presented in curves a–e). Photocurrents reported in curve e increased to 8.05 mA cm⁻². Potential was applied by a PINE AFRDE4X1 potentiostat in a conventional three-electrode configuration (Pt reference and 6-cm² Pt-mesh counter electrodes). In the flow-through study, we compared photocurrent stability with and without electrolyte renewal. Electrolyte was delivered into the cell without turbulence by gravity feed. At higher light intensities (90 mW cm⁻²), in a PEC with an alkaline



electrolyte containing optimized K₃Fe(CN)₆/K₄Fe(CN)₆ concentrations with added cyanide (as in curve e), the photocurrent remained stable for >5,000 C cm⁻² in the flow cell and for >1,000 C cm⁻² without electrolyte renewal.

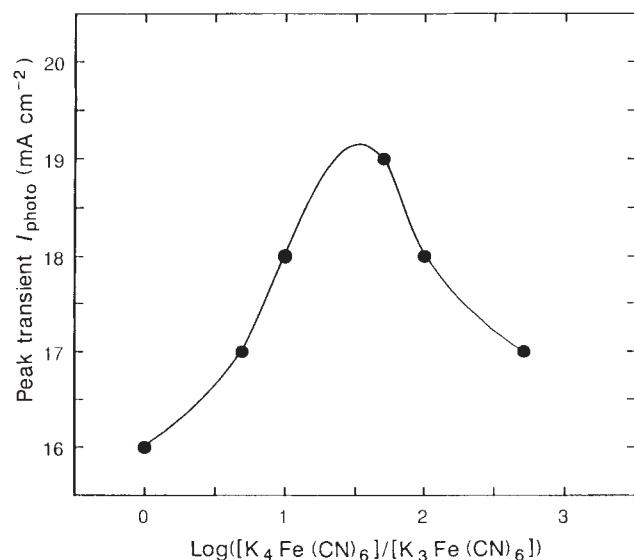
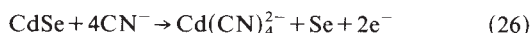


FIG. 2 Peak transient photocurrent in n-CdSe/[Fe(CN)₆^{3-/4-}]_{aq} cells measured over a wide range of K₃Fe(CN)₆ concentrations in 0.5 m K₄Fe(CN)₆ and 1.0 m KOH. The PEC was initially maintained by potentiostat at the open-circuit voltage. Peak photocurrent response was measured under a constant illumination of 25 mW cm⁻², after the application of a 300-mV anodic step. Potential steps were applied by a Heath 1274 sweep-function generator (200-ns rise time) to an AFRDE4X1 potentiostat (20-μs rise time) in a conventional three-electrode configuration (Pt reference and 10-cm² Pt-mesh counter electrodes). Current output was digitized (2 μs per data pt, 16 K data per scan) by an 80286-interfaced Heath SD-5000 Digital Scope. Electrode preparation, three-electrode cell configuration and illumination source are described in Fig. 1.

therefore expected and observed on addition of cyanide (Fig. 1, curve e). At more-intense 90 mW cm⁻² illumination, the cell remains stable for the photocurrent generation of over 5,000 C cm⁻². The polished CdSe surface can lose over 20% reflectivity²⁴. Our observed increase in photocurrent during the initial photoconversion (Fig. 1 inset, curve e) parallels an observed drop in CdSe surface reflectivity. We relate this to an initial dissolution and morphological improvement of the surface, as indicated by Cd²⁺ atomic absorption spectrophotometry (Perkin Elmer 370) of the static PEC electrolyte, coulombic count and a photoelectrode mass loss of ~0.03 g cm⁻² (equivalent to ~30 C cm⁻²). Following this initial dissolution, photocorrosion comprises ≤1% of total photocurrent, as determined by subsequent atomic absorption measurements, coulombic count and photoelectrode mass loss. Cyanide prevents passivation and chemical deposition at the CdSe surfaces, maintaining a clean surface morphology for facile charge transfer. This is in contrast with the varying degrees of success in previous attempts to modify photoelectrode surfaces (including metal², polymer³, chalcogenide⁸, electrocatalyst⁵ and mixed alkali salt⁹ surface modifiers). Thin-film n-CdSe photoelectrodes are expected to be several orders of magnitude more stable than comparable single-crystal n-CdSe PECs⁶.

There is a third beneficial effect of CN_{aq}⁻ addition. n-CdSe/KFe(CN)₆^{2-/3-} PECs containing 0.1 m KCN show open-circuit photovoltages 250 mV more negative compared with electrolytes without cyanide, and $E_{\text{redox}}\{\text{KFe(CN)}_6^{2-/3-}\}$ is not shifted (Fig. 3). Enhancement of the photovoltage is accompanied by a decrease in dark current (Fig. 3). The shift in measured open-circuit voltage (V_{oc}), ~0.12 V per decade [CN_{aq}⁻] and ~0.03 V per decade [Cd(CN)₄²⁻]_{aq} (Fig. 3 inset), is consistent with a decomposition reaction



As discussed, little CdSe decomposition is observed with passage of >5,000 C cm⁻². It is interesting that in accordance with equation (26), there is a thermodynamic shift of photo-

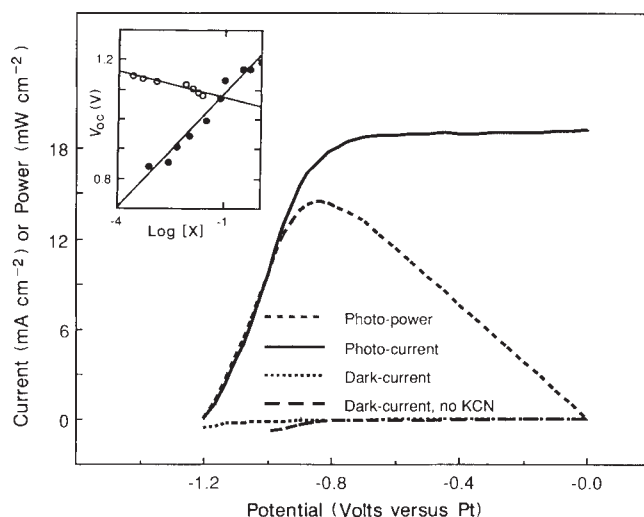


FIG. 3 Outdoor potentiostatic characteristics for an illuminated (88 mW cm⁻² insolation) or dark n-CdSe/0.25 m K₄Fe(CN)₆, 0.0125 m K₃Fe(CN)₆, 0.1 m KCN, 0.5 m KOH PEC. Electrode preparation and three-electrode cell configuration are described in Fig. 1. The results were reproduced with several n-CdSe electrodes of area 0.05–0.20 cm², at potential sweep rates of 0.5–2 mV s⁻¹. Measured solar-to-electrical conversion efficiency varied from 15.6 to 16.4%. Insolation was calibrated using an Eppley 8-48 radiometer. 'no KCN' indicates an electrolyte 0.25 m K₄Fe(CN)₆, 0.0125 m K₃Fe(CN)₆ and 0.5 m KOH. On Pt, the 0.25 m K₄Fe(CN)₆, 0.0125 m K₃Fe(CN)₆ and 0.5 m KOH electrolytes with and without 0.1 m KCN, had respective potentials of 0.132 and 0.133 V relative to the AgCl reference electrode. Inset: The variation of open-circuit photovoltage (under 80 mW cm⁻² tungsten-halogen illumination) in n-CdSe/0.25 m K₄Fe(CN)₆, 0.0125 m K₃Fe(CN)₆, 0.5 m KOH PECs containing varying concentrations of CN⁻ (filled circles), added as KCN (no added Cd(CN)₄²⁻), and in constant 0.1 m KCN with varying concentrations of Cd(CN)₄²⁻ (open circles), added as 4:1 [KCN]:[CdCl₂].

induced oxidation equation (24), without the faradaic losses suggested by equation (26). Electrochemical (dark) oxidation of CN⁻ on Pt electrodes is insignificant at these potentials and concentrations of KCN. An analysis of the adsorption/flat-band potential shift, consistent with our model for ion/semiconductor interactions²⁵, is in progress. n-CdSe chalcogenide photovoltages of 1.0 V have been previously observed in dithiocarbamate solutions and attributed to strong specific adsorption²⁶. Photovoltages in n-GaAs PECs in excess of those predicted for analogous solid-state junctions have also been observed²⁷.

Outdoors, the n-CdSe/KFe(CN)₆^{2-/3-} solar cell with an alkaline electrolyte containing optimized K₃Fe(CN)₆/K₄Fe(CN)₆ electrolyte and added KCN shows a stable conversion efficiency >16% (fill factor of 0.63) and an open-circuit photovoltage of 1.20 V (Fig. 3). The quantum yield of the alkaline n-CdSe/Fe(CN)₆^{3-/4-} PEC is known to approach unity¹⁰. CdSe with a 1.7-eV band gap has a limiting current density of 18.45 mA cm⁻² at air mass 1.5 (84.4 mW cm⁻²) insolation²⁸. Our measured photocurrents, normalized to air mass 1.5, range from 17.3 to 18.2 mA cm⁻² and are consistent with small reflectance or other losses.

In a future paper we will summarize the numerical calculations of the solution equilibria, and the effects of the variation of cation and ligand M_xFe(CN)₅L (for L = CN, NO, NO₂, NH₃ and Fe(CN)₅, and for M = NH₃, Ca, Li, Na, K and Cs).

Received 9 November 1989; accepted 6 April 1990.

1. Parkinson, B. *Accounts Chem. Res.* **17**, 431–437 (1984).
2. Miller, B., Heller, A., Menezes, S. & Lewerenz, H. J. *Discuss. Faraday Soc.* **70**, 223–232 (1981).
3. Noufi, R. *Conf. Rec. IEEE Photovoltaic Spec. Conf.* **16th**, 1293–1298 (1982).
4. Licht, S. *et al. Appl. Phys. Lett.* **46**, 608–610 (1985).
5. Tufts, B. J. *et al. Nature* **326**, 861–862 (1987).
6. Licht, S. *J. phys. Chem.* **90**, 1096–1099 (1986).
7. Licht, S., Hodes, G., Tenne, R. & Manassen, J. *Nature* **326**, 863–864 (1987).

8. Reichman, J. & Russak, M. A. *J. electrochem. Soc.* **131**, 796–798 (1984).
9. Rubin, H. D., Humphrey, B. D. & Bocarsley, A. B. *Nature* **308**, 339–341 (1984).
10. Freeze, K. W. *Jr. Appl. Phys. Lett.* **40**, 275–277 (1982).
11. Rubin, H. D., Arent, D. J. & Bocarsley, A. B. *J. electrochem. Soc.* **132**, 523–524 (1985).
12. Licht, S. *Nature* **330**, 148–151 (1987).
13. Jordan, J. & Ewing, G. *J. Inorg. Chem.* **1**, 587–591 (1962).
14. Hanania, G. I. H., Irvine, D. H., Easton, W. A. & George, P. *J. phys. Chem.* **71**, 2022–2030 (1967).
15. Capone, S., de Robertis, A., de Stefano, C. & Scarella, R. *J. chem. Res.* **1**, 412–413 (1986).
16. Peter, L. M., Durr, W., Bindra, P. & Gerischer, H. *J. electroanal. Chem.* **132**, 31–50 (1976).
17. Lando, D., Manassen, J., Hodes, G. & Cahen, D. *J. Am. chem. Soc.* **101**, 3969–3971 (1979).
18. Licht, S., Tenne, R., Flaisher, H. & Manassen, J. *J. electrochem. Soc.* **133**, 52–59 (1985).
19. Itaya, K., Akahoshi, H. & Toshima, S. *J. electrochem. Soc.* **129**, 1498–1500 (1982).
20. Asperger, S., Murati, I. & Pavlovic, D. *J. chem. Soc.* 730–736 (1960).
21. Lal, B. B. *Ind. chem. Soc. J.* **16**, 321–323 (1939).
22. Ayers, J. B. & Waggoner, W. H. *J. inorg. nucl. Chem.* **33**, 721–733 (1971).
23. Reihlen, H. & Zimmermann, W. *Justus Liebigs Ann. Chem.* **475**, 101–119 (1929).
24. Heller, A., Chang, K.-C. & Miller, B. *J. Am. chem. Soc.* **100**, 684–688 (1978).
25. Licht, S. & Marcu, V. *J. electroanal. Chem.* **210**, 197–204 (1986).
26. Thackeray, J. W., Natan, M. J., Ng, P. & Wrighton, M. S. *J. Am. chem. Soc.* **108**, 3570–3577 (1986).
27. Abrahams, I. L. *et al. New J. Chim.* **11**, 157–165 (1987).
28. Henry, C. H. *J. appl. Phys.* **51**, 4494–4500 (1980).

ACKNOWLEDGEMENTS. S.L. is grateful to the Carl Julius and Anna (Kranz) Carlson Chair in Chemistry for financial support.

Heterogeneous reactions on nitric acid trihydrate

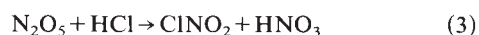
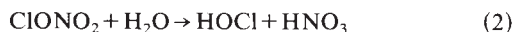
S. B. Moore*, L. F. Keyser*, M.-T. Leu*,
R. P. Turco† & R. H. Smith*‡

* Earth and Space Sciences Division, Jet Propulsion Laboratory,
California Institute of Technology, Pasadena,
California 91109, USA

† Department of Atmospheric Sciences, University of California,
Los Angeles, California 90024, USA

HETEROGENEOUS reactions on the surfaces of polar stratospheric cloud particles^{1–4} are now known to have a crucial role in the springtime decrease of polar ozone^{5,6} by converting inactive chlorine reservoir species, ClONO₂ and HCl, to photochemically active forms, Cl₂ and HOCl, thus preparing the stratosphere for ozone destruction on the return of sunlight in the spring. There is abundant recent theoretical^{7–9} and indirect observational^{10–16} evidence for two main types of polar stratospheric cloud (PSC). The predominant form, type I, which comprises HNO₃ and H₂O in the form of frozen nitric acid trihydrate (NAT), can condense at 195 K in the stratosphere⁹. This is 5–7 K warmer than the condensation temperature of type II PSCs, which are composed mainly of water ice. To understand the surface chemistry of PSCs and to test current hypotheses concerning polar ozone loss, laboratory measurements of several key kinetic parameters are needed. Here we report the first direct measurements of the reaction probabilities at stratospheric temperatures for two important heterogeneous reactions on NAT. Additionally, we report the sticking coefficients and solubilities of HCl and NAT, which are important in modelling physical-chemical processes in the stratosphere. Our results show that the conversion of the chlorine reservoir species can occur within a few days of the first appearance of type I PSCs during the polar winter.

The heterogeneously catalysed reactions of primary interest in the stratosphere are



Reaction (1) may be particularly important as it converts the chlorine in two inactive reservoir molecules into the volatile and photochemically active Cl₂ molecule. Reaction rates have been measured for reactions (1)–(4) on ice^{17–21} but not on NAT.

Rates of reactions (1) and (2) were measured by using a fast-flow reactor coupled to a differentially pumped mass-spectrometer detection system (Fig. 1)^{20,21}. Total pressures were ~0.4–2 torr at a temperature of 196 K; average flow velocities ranged from ~100 to 2,500 cm s⁻¹. Substrates were formed on the inner surface of the cooled reactor tube by vapour deposition at controlled rates near 0.5 g h⁻¹. Compositions were varied from 37–57 weight per cent (wt%) HNO₃ to cover the range of NAT compositions⁹ (45–53 wt% HNO₃) believed to be stable in the stratosphere. As used here, the term NAT includes both H₂O-rich and HNO₃-rich ices near the composition of pure nitric acid trihydrate, 53.83 wt% HNO₃. The HNO₃-ice substrates seemed to be frost-like crystalline solids and served as approximations of actual PSC aerosols.

In separate experiments, we also investigated the chemical composition of NAT substrates by means of Fourier transform infrared spectroscopy. Thin (~1 μm) substrates were deposited very slowly on an AgCl window which was maintained at 188 K. The infrared spectra were collected while the vapour deposition took place or after annealing the substrates for about an hour. The characteristic infrared absorption spectrum of NAT was observed, and the spectra indicate that the NAT substrates may contain hydronium and nitrate ions.

To measure the rate of reaction (1), the substrate is first doped with HCl by vapour deposition and then ClONO₂ is continuously admitted to the reactor through the injector (see Fig. 1). Figure 2 shows the change in the reactor-exit concentrations of the ClONO₂ reactant and Cl₂, the product of reaction (1),

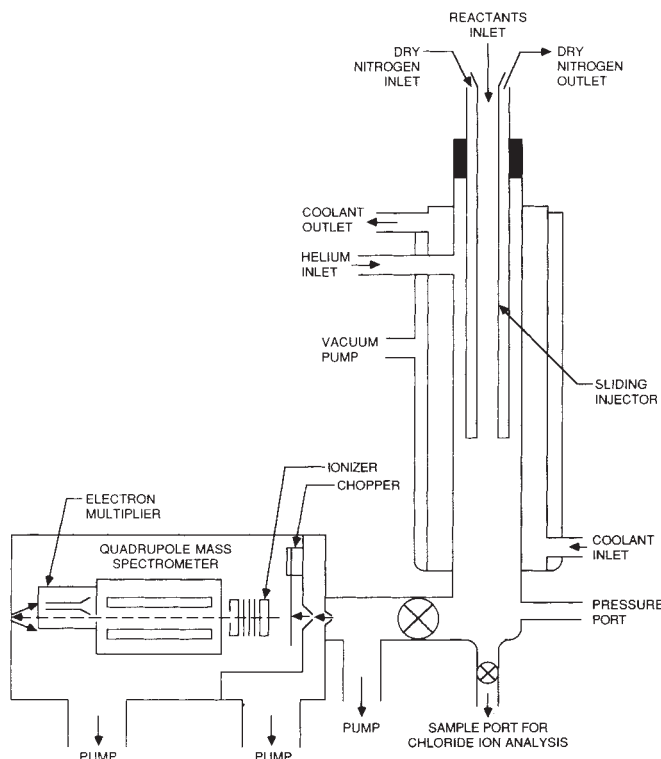


FIG. 1 Schematic diagram of fast-flow reactor system. A thin (~50 μm) NAT substrate is formed on the internal surface of the cylindrical reactor (1.76 cm in diameter) for a distance of 30 cm starting at the reactor exit by flowing He carrier gas saturated with HNO₃ and H₂O (at 293 K) through the injector after the reactor is cooled to operating temperature (196 ± 1 K). Usually, ~1 g of substrate is deposited. HCl is deposited in the substrate by flowing a mixture of He and HCl over the surface, in some cases saturating the substrate with HCl. Subsequently, He carrier gas at 0.4–2.0 torr is used to flow the reactants into the reactor. The HCl flows in through the main He inlet and the ClONO₂ flows in through the injector to prevent reaction except over the substrate. Following an experiment, the substrate is warmed and collected for analysis.

‡ Permanent address: School of Chemistry, MacQuarie University, North Ryde, New South Wales 2113, Australia.

as the contact time over the substrate is increased by moving the injector. The observed loss rate of ClONO_2 and the production rate of Cl_2 are used to obtain a first-order surface rate constant k_s by standard kinetics analysis techniques^{20,21}. The parameter of interest in atmospheric modelling is the reaction probability γ which can be obtained from k_s by equating the reaction rate with the surface collision rate which depends on the total surface area; for a cylindrical reactor, this leads to the relation

$$\gamma = 2ak_s / (\omega + ak_s) \quad (5)$$

where a is the reactor radius and ω is the molecular velocity of ClONO_2 . Before using the above equation to calculate γ , it is necessary to correct the observed surface rate constants for the interaction of flow dynamics, diffusion and reaction²². Under our experimental conditions, corrections were less than 20% for $\gamma < 0.01$; but for $\gamma > 0.1$ the corrections were a factor of two or larger.

The observed reaction probabilities for ClONO_2 with HCl are shown in Fig. 3 as a function of the HCl concentration in the substrate. Values of γ range from 0.06 to 1.0 with an overall average of 0.3. Our observation of uniformly high reaction probabilities for all HCl concentrations used may indicate that the surface is saturated with HCl at the partial pressures,

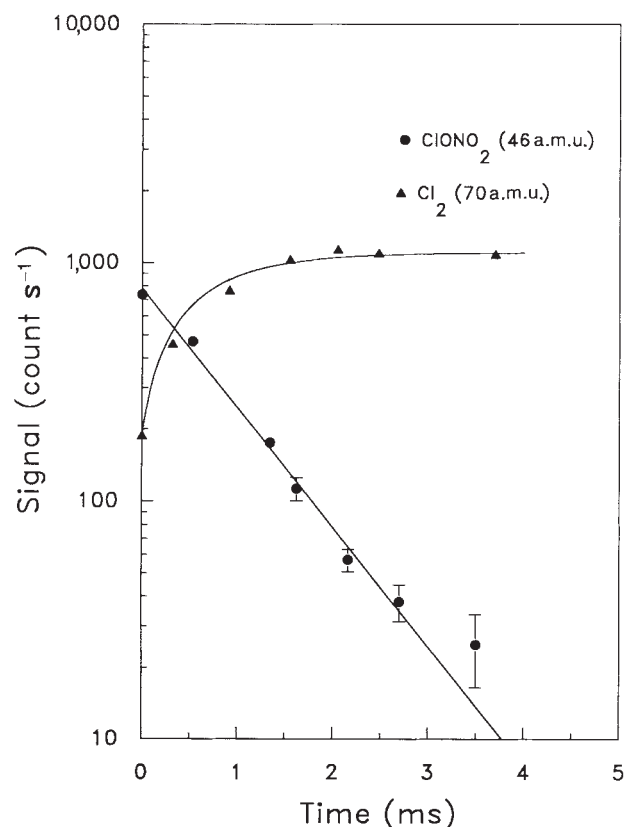


FIG. 2 Typical raw mass spectrometer signal, which is proportional to the reactor-exit concentration, for the $\text{ClONO}_2 + \text{HCl} \rightarrow \text{Cl}_2 + \text{HNO}_3$ reaction on a NAT substrate. The length of substrate exposed to ClONO_2 has been converted to time by dividing by the average flow velocity (1.856 cm s^{-1}). The NO_2^+ ion fragment from ClONO_2 is monitored with the mass spectrometer, and Cl_2 is monitored by its parent ion, Cl_2^+ . The linear decay of the ClONO_2 signal on the logarithmic plot implies the reaction is first order with respect to ClONO_2 ; the slope of the line gives the first-order rate constant for reaction (1). The deviation of the ClONO_2 signal from linearity at long times is the result of interference from HNO_3 evaporating from the substrate; the HNO_3 produced in the reaction is incorporated in the substrate and not released to the gas phase immediately. The data are from successive runs on a substrate comprised of 53.2 wt% HNO_3 doped with 0.46 wt% HCl at a temperature of 196 K.

$P(\text{HCl})$, used in these experiments (0.1–0.3 mtorr). We observed no significant change in γ between successive runs on fresh substrate in which the $P(\text{HCl})$ was either held constant by a continuous flow of HCl into the reactor or allowed to drop to a much lower value (presumably its equilibrium level over the substrate). This may indicate that the surface remains saturated with HCl at these lower partial pressures or that the HCl is rapidly replenished at the surface from the interior. It is possible that at the much lower $P(\text{HCl})$ s typical of the stratosphere ($< 10^{-6}$ torr) the NAT surfaces will not be saturated with HCl and γ will be correspondingly lower and a function of $P(\text{HCl})$. The observed reaction probabilities are independent of the HNO_3 substrate concentrations over the range studied (42–53.6 wt% HNO_3) on the water-rich side of pure NAT. At HNO_3 substrate concentrations greater than 54 wt%, HCl sorption decreases rapidly (see below). We were not able to maintain a constant HCl substrate concentration, and therefore could not carry out measurements of reaction (1) on acid-rich substrates. The good agreement between reaction probabilities determined from monitoring the ClONO_2 reactant concentration and these determined from monitoring the Cl_2 product concentrations (see Fig. 3) indicates that a true reaction probability was being measured, as opposed to the physical adsorption of ClONO_2 . These observations are consistent with the results for water ice reported previously from our laboratory²⁰.

The γ of ClONO_2 on NAT in the absence of HCl was determined to be at least an order of magnitude smaller than that in the presence of HCl (0.02 ± 0.01 at 45–47 wt% HNO_3) and to fall sharply as the composition of pure NAT was approached. Measurements of the physical sticking coefficient of HCl on NAT substrates were carried out by using the same technique used to determine γ . The results showed similar

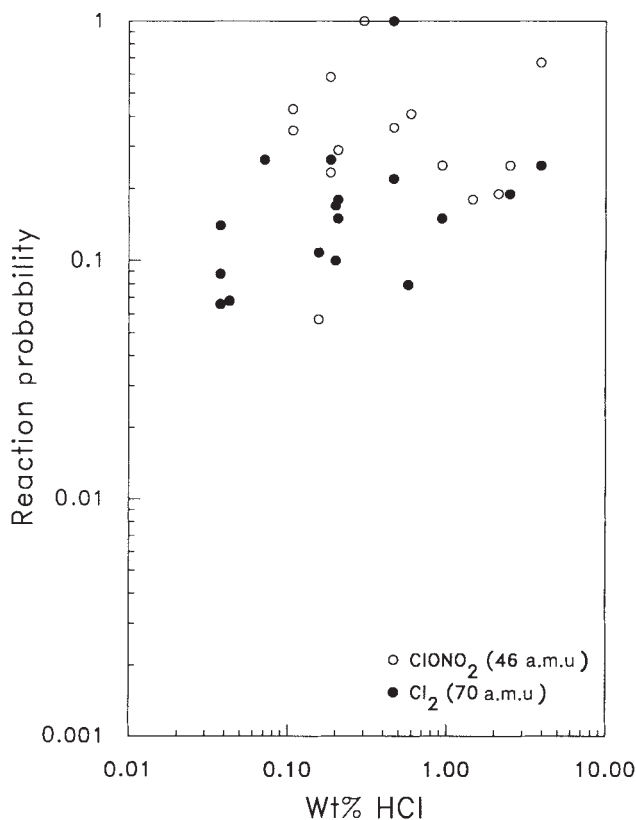


FIG. 3 The reaction probability γ of ClONO_2 with HCl on NAT substrate. The substrate HNO_3 composition varies from 42.0 wt% to 53.6 wt%. Reaction probabilities are obtained both from the decay of ClONO_2 and the growth of Cl_2 . A correction for gas-phase diffusion in the reactor has been incorporated in the calculated reaction probabilities.

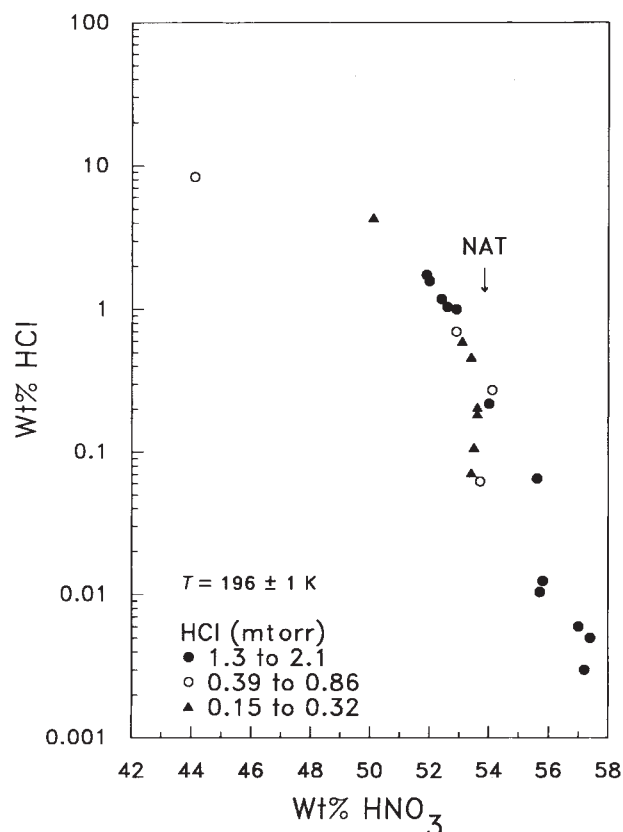


FIG. 4 Sorption of HCl by NAT as a function of wt% HNO_3 . The arrow indicates the composition of pure NAT, 53.83 wt%. We calibrated the mass spectrometer with a given partial pressure of HCl (monitored with HCl^+) in the He flow. Equilibrium was assumed to occur when the signal for the HCl flowing over the substrate returned to the calibration level. The total acid content of the substrate is determined by standard acid-base titration, and the HCl content by means of a Cl^- selective electrode.

behaviour, falling from 0.05 ± 0.01 at 44–46 wt% HNO_3 to less than 10^{-4} at compositions greater than 55 wt% HNO_3 .

The calculation of a reaction probability or a sticking coefficient requires a knowledge of the effective surface area of the substrate. The values reported above are based on the macroscopic external surface area of the substrate. Brunauer–Emmett–Teller (BET) measurements of the substrate's microscopic surface area give values ~ 10 – 100 times larger than the macroscopic area. The amount of this area that is available to the reacting species depends on the exact structure of the substrate. A model to take this effect into account is currently being developed. We estimate that reaction probabilities greater than 0.1 would be reduced by about a factor of 5, but measured probabilities less than 0.01 could be significantly reduced.

The substrates were also saturated with HCl to determine the

equilibrium sorption of HCl by the NAT substrate. The results (Fig. 4) are in qualitative agreement with solubility measurements^{23,24} made at much lower $P(\text{HCl})$ s (10^{-8} – 10^{-5} torr), which required temperatures near the frost point (that is, excess water of hydration), before measurable amounts of HCl were taken up by NAT. We were not able to determine the mode of HCl incorporation in the NAT, and it may exist as a highly concentrated phase on the surface or at grain boundaries²⁵. It is clear that more than simple physical adsorption is involved because a monolayer (based on a molecular area of $16 \times 10^{-20} \text{ m}^2 \text{ molecule}^{-1}$) of HCl condensed on a $3 \text{ m}^2 \text{ g}^{-1}$ solid gives ~ 0.1 wt% HCl—far less than our observation of at least 0.7 wt% HCl for substrates with less than 53 wt% HNO_3 .

We may now consider the implications of our measurements for theories of the ozone-hole formation. For reaction (1), the present study gives a reaction probability, $\gamma \approx 0.1$ – 0.3 , on the HCl-doped water-enriched type I ices which are expected under stratospheric conditions^{15,23}; for reaction 2, $\gamma \approx 0.02$. Hence, HCl and ClONO_2 conversion to Cl_2 through reaction (1) can occur within one day and ClONO_2 activation through reaction (2) within one week²⁶. The initial conversion of the chlorine reservoir thus occurs within a few days of the first appearance of type I PSCs in the polar winter season, and the chlorine reservoir remains activated as long as these nitric acid PSCs continue to form or NO_2 is absent (through denitrification). Our data suggest that, as type I PSCs are also common in the Arctic polar winter stratosphere^{13,27,28}, chlorine activation will also occur in the north.

The earlier work^{23,24} and the present results show that type I cloud particles probably contain only ~ 0.5 wt% of HCl compared to ~ 50 wt% of HNO_3 whereas the ambient stratospheric ratio of HCl: HNO_3 is roughly 1:10. Thus, the chlorine reservoir cannot be significantly depleted by sedimentation of type I cloud particles (concurrent with denitrification).

The close coupling between type I PSCs and chlorine activation provides a mechanism for prolonging the ozone hole during the austral spring. The southern polar vortex is destabilized in October/November by warming of the vortex associated with the absorption of sunlight by ozone. If ozone is strongly depleted in October, the warming is reduced and the vortex remains cold and stable. PSCs are thus more persistent²⁹, chlorine continues to be activated³⁰, and the depleted levels of ozone can persist. The possibility of such a positive chemical/radiative/dynamical feedback mechanism is in part the result of heterogeneous chlorine activation on the particles of nitric acid ice, as defined by the present study.

Farman *et al.*⁵ first attributed the development of the ozone hole to the accumulation of chlorofluorocarbons (CFCs) in the atmosphere. The chlorine released when CFCs are photochemically decomposed in the stratosphere can be activated on PSCs and deplete polar ozone in the spring^{1–3}. Our results indicate that the observed onset of the ozone depletion is indeed connected to the buildup of CFCs, and that ozone depletion has probably never been strongly limited by chlorine activation on PSCs. □

Received 23 January; accepted 3 April 1990.

- Solomon, S., Garcia, R. R., Rowland, F. S. & Wuebbles, D. J. *Nature* **321**, 755–758 (1986).
- McElroy, M. B., Salawitch, R. J., Wofsy, S. C. & Logan, J. A. *Nature* **321**, 759–762 (1986).
- Crutzen, P. J. & Arnold, F. *Nature* **324**, 651–655 (1986).
- Hofmann, D. J., Harder, J. W., Rolf, S. R. & Rosen, J. M. *Nature* **326**, 59–62 (1987).
- Farman, J. C., Gardiner, B. G. & Shanklin, J. D. *Nature* **315**, 207–210 (1985).
- Stolarski, R. S. *et al.* *Nature* **322**, 808–811 (1986).
- Toon, O. B., Hamill, P., Turco, R. P. & Pinto, J. *Geophys. Res. Lett.* **13**, 1284–1287 (1986).
- McElroy, M. B., Salawitch, R. J. & Wofsy, S. C. *Geophys. Res. Lett.* **13**, 1296–1299 (1986).
- Hanson, D. & Mauersberger, K. *Geophys. Res. Lett.* **15**, 855–858 (1988).
- Austin, J., Remsberg, E. E., Jones, R. L. & Tuck, A. F. *Geophys. Res. Lett.* **13**, 1256–1259 (1986).
- Poole, L. R. & McCormick, M. P. *Geophys. Res. Lett.* **15**, 21–23 (1988).
- Hofmann, D. J., Rosen, J. M., Harder, J. W. & Hereford, J. V. *J. geophys. Res.* **94**, 11,253–11,269 (1989).
- Rosen, J. M., Oltams, S. J. & Evans, W. F. *Geophys. Res. Lett.* **16**, 791–794 (1989).
- Pueschel, R. F. *et al.* *J. geophys. Res.* **94**, 11,271–11,284 (1989).
- Fahey, D. W. *et al.* *J. geophys. Res.* **94**, 11,299–11,315 (1989).
- Arnold, F., Schlager, H., Hoffmann, J., Metzinger, P. & Spreng, S. *Nature* **342**, 493–497 (1989).

- Tolbert, M. A., Rossi, M. J., Malhotra, R. & Golden, D. M. *Science* **238**, 1258–1260 (1987).
- Tolbert, M. A., Rossi, M. J. & Golden, D. M. *Science* **240**, 1018–1021 (1988).
- Molina, M. J., Tso, T. L., Molina, L. T. & Wang, F. C.-Y. *Science* **238**, 1253–1257 (1987).
- Leu, M.-T. *Geophys. Res. Lett.* **15**, 17–20 (1988).
- Leu, M.-T. *Geophys. Res. Lett.* **15**, 851–854 (1988).
- Brown, R. L. *J. Res. natn. Bur. Stand.* **83**, 1–8 (1978).
- Hanson, D. & Mauersberger, K. *Geophys. Res. Lett.* **15**, 1507–1510 (1988).
- Hanson, D. & Mauersberger, K. *J. phys. Chem.* (in the press).
- Wolff, E. W., Mulvaney, R. & Oates, K. *Geophys. Res. Lett.* **16**, 487–490 (1989).
- Turco, R. P., Toon, O. B. & Hamill, P. *J. geophys. Res.* **94**, 16,493–16,510 (1989).
- McCormick, M. P. & Trepte, C. R. *J. geophys. Res.* **92**, 4297–4306 (1987).
- Hamill, P., Turco, R. P. & Toon, O. B. *J. Atmos. Chem.* **7**, 287–315 (1988).
- McCormick, M. P., Trepte, C. R. & Pitts, M. C. *J. geophys. Res.* **94**, 11,241–11,251 (1989).
- Poole, L. R., Solomon, S., McCormick, M. P. & Pitts, M. C. *Geophys. Res. Lett.* **16**, 1157–1160 (1989).

ACKNOWLEDGEMENTS. This research was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA.

Geometry of subduction and depth of the seismogenic zone in the Guerrero gap, Mexico

Gerardo Suárez*, Tony Monfret*†, Gerard Wittlinger‡ & Christian David‡

* Instituto de Geofísica, Universidad Nacional Autónoma de México, México DF 04510, México

† Mission ORSTOM-Mexique, Homero 1804-1002 México DF 11510, México

‡ Institut de Physique du Globe, 5 rue René Descartes, Strasbourg 68084 Cedex, France

THE western coast of Guerrero in southern Mexico has been identified as a seismic gap on the Middle American Trench in which no large earthquakes have occurred at least since 1908. It has been suggested that the seismic energy accumulated since 1908 will eventually be released by a large earthquake. A permanent seismic network was installed to monitor the seismicity of this mature seismic gap and to understand the geometry of the subducted slab beneath this region. The seismicity defines an unusual distribution along two bands of seismic activity parallel to the coast. The resulting geometry of the subduction zone shows that the Cocos plate dips at a shallow angle beneath the North American plate to a depth of ~ 40 km; from there the subducted slab is bent upward, following a subhorizontal trajectory extending inland at a depth of ~ 50 km. This plate geometry is reminiscent of that found in Peru and central Argentina, two other regions where a young oceanic plate is being subducted. In Mexico, however, the slab underplates an overriding plate which is only half as thick as that observed in South America. A possible oceanic origin of the allochthonous terranes comprising southern Mexico may explain the presence of the anomalously thin lithosphere in this region.

The Guerrero gap in southern Mexico is perhaps one of the more clearly identified seismic gaps in the circum-Pacific belt¹⁻³. It lies immediately south of the rupture area of the 1985 Michoacan earthquake^{4,5} (Fig. 1). A telemetered, nine-station seismic network northwest of Acapulco (Fig. 1) records daily an average of four to five earthquakes, with coda-wave magnitude (M_c) in the range 1-4, which are located using the program HYPO71 (ref. 6) and a velocity model reported for this area⁷. The results presented here include events recorded during a three-month temporary experiment performed in 1986 and earthquakes located from August 1987 to December 1988 by the permanent network.

The distribution of seismicity shows an unusual disposition along two bands of activity (Fig. 1). The coastal band of seismicity is ~ 35 km wide and shows hypocentres with focal depths of between 10 and 25 km. The second zone of seismicity lies farther inland and is clearly separated from the coastal activity, showing focal depths of between 32 and 42 km. The absence of earthquakes between these two seismic bands is more evident where the station coverage is best. Also, practically no seismic activity is located between the coastal seismic zone and the trench (Fig. 1). Although this area is outside the network, the absence of seismic activity here is not due to its distance to the seismic instruments on land; small earthquakes ($0 < M_c < 1$) are routinely located along the coast at equivalent distances from the centre of the network. This low level of near-trench seismicity appears where saturated sediments may have a dominant role in the mechanical behaviour of the plate interface.

In cross-section, the hypocentral distribution reflects the subduction of the Cocos plate beneath Mexico (Fig. 2). The void of seismicity between depths of 25 and 32 km, separating the two parallel seismic zones, is clearly visible on cross-sections where the accuracy of the locations is better (BB', CC' and DD'). The apparent absence of this void in section AA' probably reflects poor hypocentral control of earthquakes located outside

FIG. 1 a, Map showing the location of stations of the Guerrero seismic network (solid triangles) and of the selected epicentres. Cross-sections are shown in Fig. 2. Composite fault-plane solutions are in lower-hemispheric projection; solid quadrants indicate compressional first motions. b, Location of the Guerrero gap relative to the rupture areas (ovals) of previous large earthquakes in the region. Station coordinates: PDE, 17.46°N, 100.74°W; TET, 17.16°N, 100.63°W; NUX, 17.21°N, 100.75°W; PAP, 17.30°N, 101.04°W; POP, 17.02°N, 100.24°W; FLO, 17.22°N, 100.39°W; POG, 17.37°N, 100.62°W; PDG, 17.47°N, 100.18°W; SJR, 17.14°N, 100.47°W.

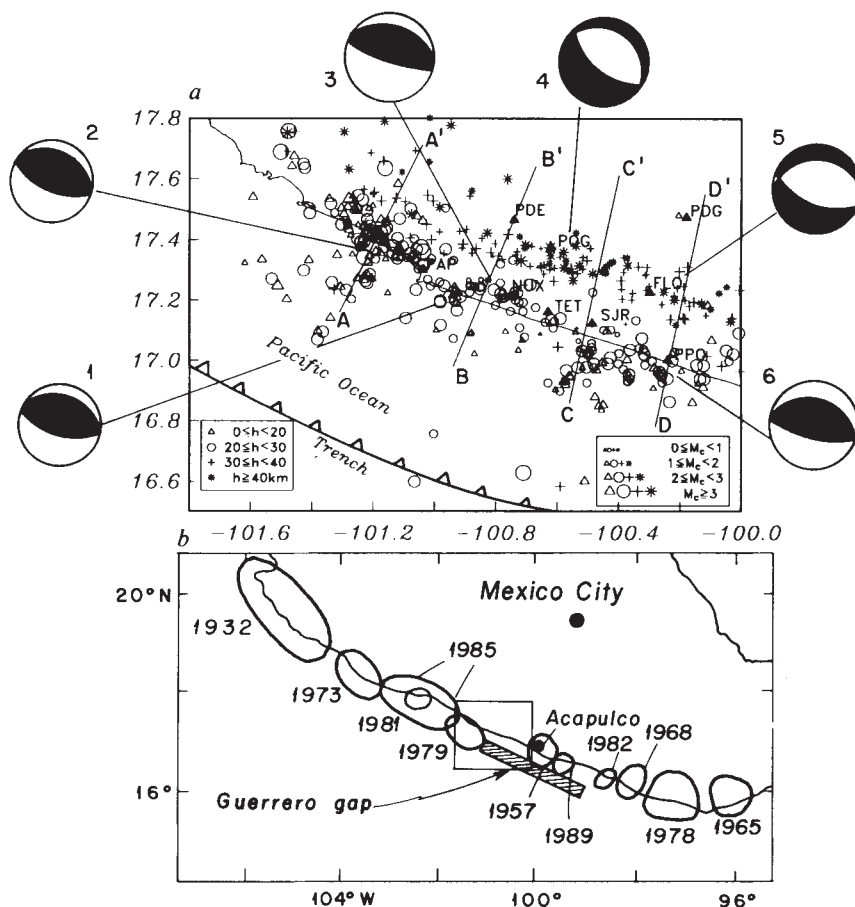
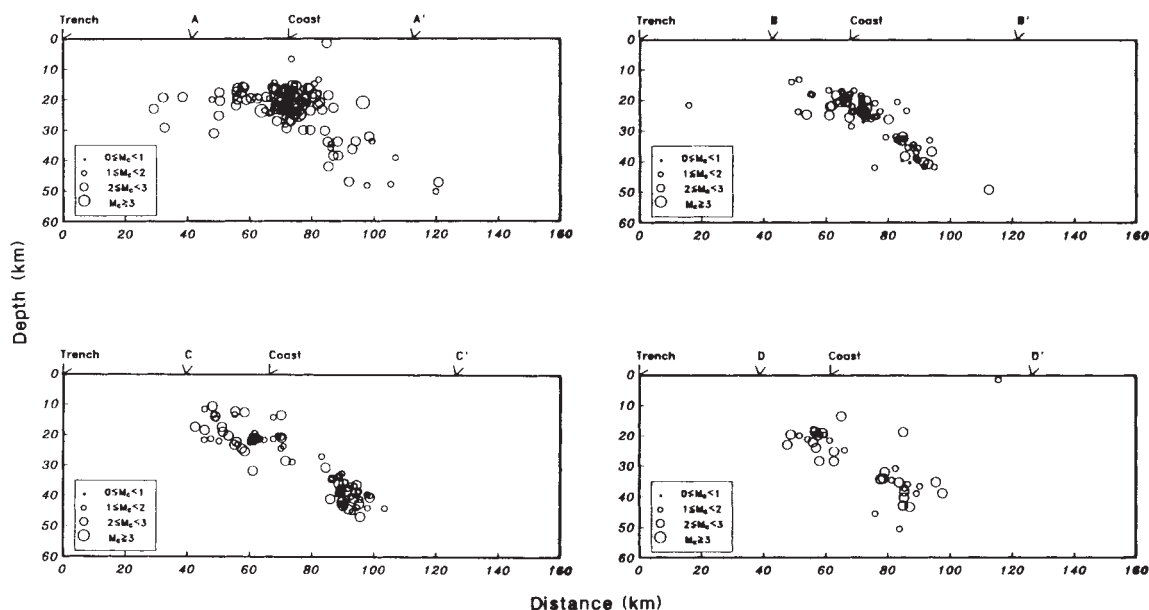


FIG. 2 Cross-sections of the seismicity (open circles) shown on Fig. 1. Size of symbols indicate magnitude of hypocentres. Notice the two distinct seismic clusters in the three cross-sections made within the seismic network (BB', CC' and DD').



of the network. Notice also that the overriding plate is virtually aseismic (Fig. 2). Composite fault-plane solutions of the coastal earthquakes show shallow-thrust faulting dipping to the north-east, reflecting the relative motion of the Cocos and North American plates (Figs 1 and 3). On the other hand, the deeper seismic zone shows predominantly normal faulting with tensional axes almost horizontal and oriented in the direction of relative plate motion (Figs 1 and 3).

No microseismicity was located inland of the second seismic band. Because of this, the trajectory of the subducted slab farther inland can not be determined only from the microearthquakes located by the network. Nevertheless, some large events occurring inland have been studied using teleseismic data, such as the Tlapachula event of 6 June 1964 ($M_s = 6.7$). Its focal mechanism and hypocentral depth are well constrained using body-wave modelling⁸, indicating normal faulting with almost horizontal T-axis oriented NNE-SSW (Fig. 4) at a depth of 55 km, similar to that of other tensional earthquakes within the subducted slab in central Mexico⁹⁻¹³. Another earthquake on 2 July 1968 also shows (ref. 14) normal faulting at a depth of 45 km¹⁴ in the region where no seismicity has been recorded by the network. The geometry of the slab is inferred by projecting the hypocentre of these two earthquakes onto a cross-section together with the microseismic data (Fig. 4).

The data show that the Cocos plate subducts beneath the North American plate at a shallow angle which steepens progressively to $\sim 12^\circ$. The maximum depth of strong seismogenic plate contact is 25 km. Beneath this depth the subducted slab bends sharply, following a quasi-horizontal trajectory underplating southern Mexico for at least 150 km (Fig. 4). The narrow coastal seismicity reflects the relative motion of the Cocos and North American plates, whereas the tensional earthquakes located deeper and farther inland seem to reflect flexural stresses induced by the sharp bend in the slab. The morphology of the slab to the north of the flat portion is still not known (Fig. 4). Probably the slab steepens north of this region, reaching a depth of ~ 100 km beneath the volcanic belt, as observed in most subduction zones¹⁵. The apparent buoyancy of the slab in central Mexico may be due to the young age of the Cocos plate. In southeastern Mexico and Central America, where the Cocos plate is older¹⁶, the slab dip becomes steeper (30°) with no evidence of a horizontal segment¹⁰⁻¹².

The flexure of the subducted slab near the coast may explain also the presence of great, lithospheric tensional earthquakes occurring in Mexico immediately downdip of the megathrust plate contact. The 15 January 1931 event in Oaxaca ($M_w = 8.0$)¹³, and probably the great earthquake of 19 June 1858 in northern

Michoacan ($M_w \approx 8.0$), are difficult to explain by the purported gravitational pull of a slab that is subhorizontal and shallower than ~ 100 km in central Mexico^{11,12}. Their origin is perhaps better explained by flexural stresses present in the contorted slab.

This bend in the slab at a depth of ~ 30 km apparently limits

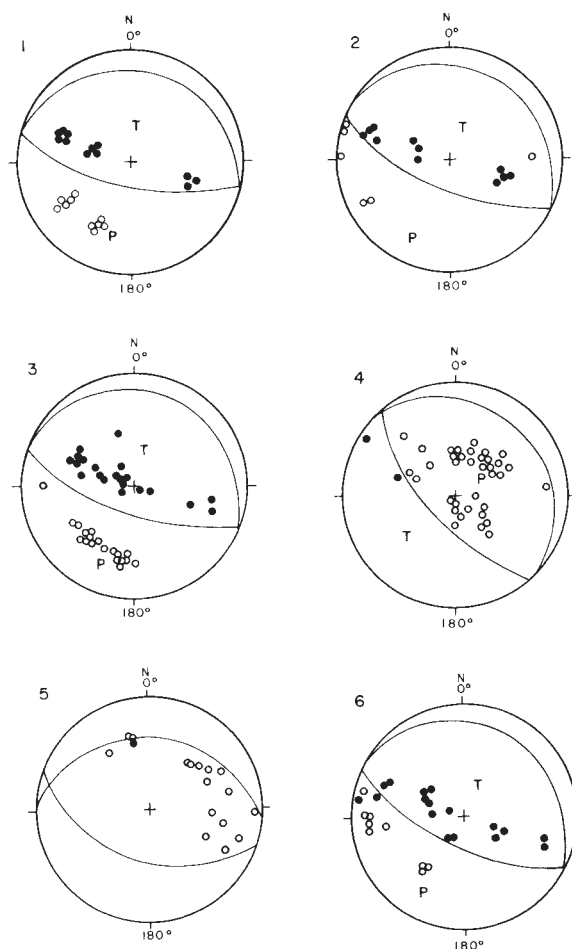


FIG. 3 Composite focal mechanisms of the earthquakes shown on Fig. 1. Solid circles indicate compressional first motions; smaller symbols are indicative of nodal arrivals. The letters P and T indicate axes of maximum and minimum compression respectively. Projection is on a lower-hemisphere stereonet.

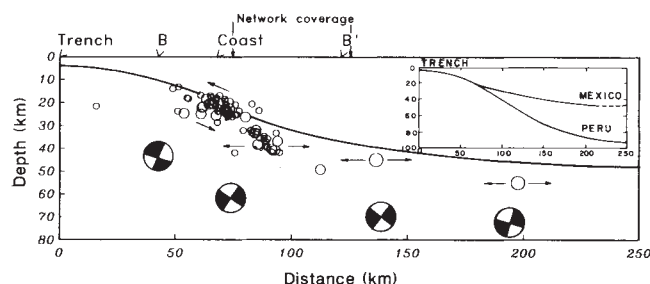


FIG. 4 Geometry of the subducted slab beneath southern Mexico. Seismicity corresponds to that of cross-section CC' on Figs 1 and 2. Focal mechanisms of microearthquakes and of teleseismic events^{8,14} are shown on side-looking, lower-hemisphere projection. Dark quadrants indicate compressional first motions. Arrows indicate approximate projected orientation of the T-axes. Insert shows comparison between subduction zones in central Peru^{22,23} and southern Mexico. Notice the continental lithosphere in Peru is twice as thick as that of Mexico.

the downdip extent of the seismogenic zone, explaining the unusually narrow megathrust plate contact found in the Mexican subduction zone; a fact corroborated by aftershock studies performed after large subduction earthquakes in Mexico^{4,5,7,17-19}, which show that the rupture areas of interplate events in this region are consistently shallower than 20 km. This relatively narrow seismogenic interface in Mexico may explain why the Middle American arc, where young oceanic lithosphere is being subducted at a fast rate (~ 6.5 cm yr⁻¹), does not have the large contact zone that results in major earthquakes ($M_w > 9.0$) like those of 1960 in Chile or 1964 in Alaska. The seismogenic zone in Guerrero is as narrow as the 1985 Michoacan rupture. Thus, the length of the northwestern Guerrero gap would result in an earthquake of the same magnitude as the 1985 event ($M_w = 8.1$). If the rupture propagates south of Acapulco, the magnitude could be as high as 8.4.

The plate geometry in Guerrero is reminiscent of that observed in central Peru²⁰⁻²³ and Argentina^{12,20,24}, where the Nazca plate lies subhorizontally, underplating the South American plate. central Peru, Argentina, and central Mexico are apparently the only three subduction zones that exhibit a horizontal slab beneath the upper continental plate. However, the main difference between the subduction in Mexico and those of South America is the lithospheric thickness of the overriding plate. Locally recorded microearthquakes^{21-23,25} and accurate depth determinations of teleseismic events^{22,24}, show the horizontal seismic zone beneath central Peru and Argentina lies at an average depth of 100 km, which corresponds to a lithospheric thickness of ~ 90 km. In contrast, the average lithospheric thickness in southern Mexico is only ~ 45 km (Fig. 4).

This thin lithosphere in southern Mexico is more similar to that of an oceanic plate²⁶ than to the thicker continental lithosphere²⁷ expected in this region. This is probably explained by the fact that southern Mexico was formed by the accretion of allochthonous terranes onto the continent since the Palaeozoic²⁸⁻³¹. The thin lithosphere underlying these allochthonous terranes suggests they may be oceanic in nature. Otherwise, the lithosphere in southern Mexico could be originally continental and later thinned by an extensional phase of deformation when swarms of mafic dykes were intruded^{29,32}. In any case, the young and presumably buoyant subducted slab seems simply to follow the lower topography of this thin overriding plate. □

- Stolte, C. *et al.* *Geophys. Res. Lett.* **13**, 577-580 (1986).
- Lee, W. H. K. & Lahr, J. C. *U.S. Geol. Surv., Open File Report* 75-311 (1975).
- Valdes, C., Meyer, R. P., Zúñiga, R., Havskov, J. & Singh, S. K. *J. geophys. Res.* **87**, 8519-8527 (1982).
- González-Ruiz, J. thesis, Univ. of California, Santa Cruz (1986).
- Jiménez, Z. & Ponce, L. *Geofis. Int.* **17**, 379-386 (1979).
- Nixon, G. T. *Bull. geol. Soc. Am.* **93**, 514-523 (1982).
- Burbach, G. V., Frohlich, C., Pennington, W. D. & Matsumoto, T. *J. geophys. Res.* **89**, 7719-7735 (1984).
- Bevis, M. & Isacks, B. *J. geophys. Res.* **89**, 6153-6170 (1984).
- Singh, S. K., Suárez, G. & Domínguez, T. *Nature* **317**, 56-58 (1985).
- Molnar, P. & Sykes, L. R. *Bull. geol. Soc. Am.* **80**, 1639-1684 (1969).
- Isacks, B. & Barazangi, M. *Maurice Ewing Series* **1**, *Am. Geophys. Union*, 99-114 (1977).
- Couch, R. & Woodcock, S. *J. geophys. Res.* **86**, 1829-1840 (1981).
- Ponce, L., McNally, K. C., González, J., Del Castillo, L. & Chael, E. *Geofis. Int.* **17**, 267-280 (1979).
- Singh, S. K. *et al.* *Science* **207**, 1211-1213 (1980).
- Havskov, J., Singh, S. K., Nava, E., Domínguez, T. & Rodríguez, M. *Bull. seism. Soc. Am.* **73**, 449-458 (1983).
- Barazangi, M. & Isacks, B. *Geophys. J. R. astr. Soc.* **57**, 537-555 (1979).
- Grange, F. *et al.* *J. geophys. Res.* **89**, 6139-6152 (1984).
- Hasegawa, A. & Sacks, I. S. *J. geophys. Res.* **86**, 4971-4980 (1981).
- Suárez, G. *et al.* *Geophys. J.* (in the press).
- Araujo, M. & Suárez, G. *J. geophys. Res.* (submitted).
- Grange, F. *et al.* *Geophys. Res. Lett.* **11**, 38-41 (1982).
- Forsyth, D. *Geophys. J. R. astr. Soc.* **43**, 103-162 (1975).
- Press, F. *J. geophys. Res.* **75**, 6575-6581 (1970).
- Campa, M. F. & Coney, P. *Can. J. Earth Sci.* **20**, 1040-1051 (1983).
- Ortega-Gutiérrez, F. *Geofis. Int.* **20**, 177-202 (1981).
- Pantoja-Alor, J. & Estrada-Barraza, M. *Bol. Soc. geol. Mex.* **48**, 1-15 (1986).
- Gastil, G., Krumenacher, D. & Jenschky, W. A. *Geol. Soc. Am. Map Chart Ser.*, **M24**, (1978).
- Cserna, Z. *Bol. 62*, *Inst. Geol., Univ. Nac. Auton. Mex.* (1965).

ACKNOWLEDGEMENTS. We thank G. Ball, J. A. Estrada, J. M. Holl, and S. Solis for technical support in building and maintaining the Guerrero seismic network, and F. Megard and R. Gaulon for their support. We also thank H. Lyon-Caen, F. Ortega, L. Ponce, E. Rosenbluth, J. Urrutia and R. Zúñiga for discussions. J. Domínguez and A. Araujo helped us read the seismic data. We acknowledge support from the Ministry of Foreign Affairs of France and the Mexican National Council of Science and Technology (CONACYT).

Implications of a new acoustic target strength for abundance estimates of Antarctic krill

Inigo Everson*, Jonathan L. Watkins*, Douglas G. Bone* & Kenneth G. Foote†

* British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK

† Institute of Marine Research, PO Box 1870 Nordnes, 5024 Bergen, Norway

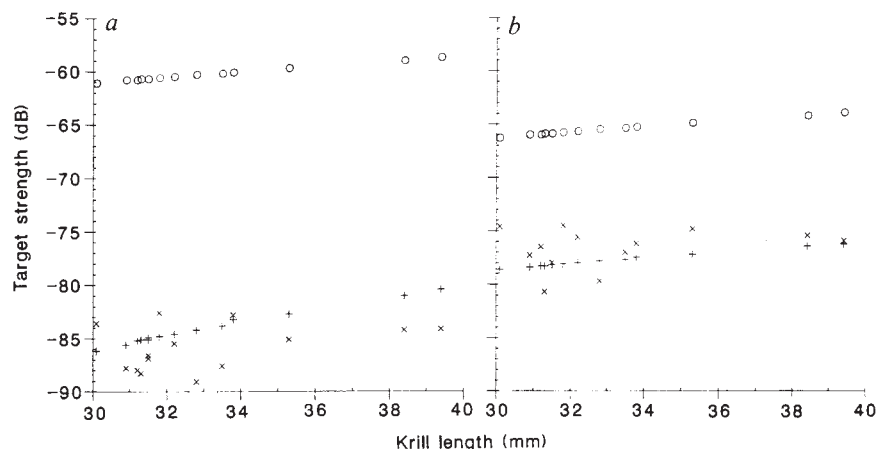
ANTARCTIC krill (*Euphausia superba*) is the dominant component of the diet of many whales, seals, birds, fish and squid, and their survival could be affected by a reduction in krill abundance due to fishing¹. Commercial fishing takes nearly half a million tonnes of krill annually² and accurate estimates of abundance are needed for rational management of this resource. Indirect estimates of abundance based on predator consumption rates give a roughly estimated total annual production of several hundred million tonnes³. The life-span of krill is at least two and maybe five years, so the standing stock would need to be of at least the same magnitude as the annual production. But direct estimates of abundance using nets and acoustics have indicated biomass figures lower by an order of magnitude³. Acoustic estimates are sensitive to the scaling factor or target strength (TS) used to convert echo energy to absolute abundance. Previous published values for TS (ref. 4), when applied to survey data, gave estimates of krill abundance that were much too low to account for local bird and seal predation rates near South Georgia⁵, and were also lower than expected when compared with density estimates from net hauls⁶. We therefore sought to determine TS using direct measurements developed for fish studies⁷, and also by applying models developed for other crustacean zooplankton⁸. Our results show that krill TS is much lower than previously thought, and consequently that acoustic estimates of krill abundance are likely to have been gross underestimates.

The field study was undertaken in Stromness Harbour, South Georgia; a comprehensive description of the methodology is published elsewhere⁹. Mean TS values for groups of krill were

Received 16 October 1989; accepted 9 March 1990.

- Kelleher, J., Sykes, L. & Oliver, J. *J. geophys. Res.* **78**, 2547-2585 (1973).
- McNally, K. & Minster, J. B. *J. geophys. Res.* **86**, 4949-4959 (1981).
- Nishenko, S. & Singh, S. K. *Bull. seism. Soc. Am.* **77**, 2095-2114 (1987).
- UNAM Seismology Group *Geophys. Res. Lett.* **13**, 573-576 (1986).

FIG. 1 A comparison of TS values at 38 kHz (a) and 120 kHz (b) derived from cage experiments (x), theoretical modelling (+) and from the BIOMASS equation (○) (ref. 4). TS in cage experiments was calculated by the subtraction method (see also Table 1). For the sound-scattering model⁸ a mean density of $1.0647 \pm 0.0069 \text{ g cm}^{-3}$ for krill of mean length 31 mm was used, and a mean sound speed contrast of 1.0279 ± 0.0024 was obtained by measuring the time delay for a sound pulse to travel through the transverse piece of an inverted T-shaped tube with ceramic transducers, resonant at 500 kHz, mounted at each end¹³. Measurements were made using clean sea water and then with increasing numbers of krill until the T-tube was more or less full. Previously used values of TS are derived from BIOMASS⁴. At 38 kHz the mean value for the model is 2.5 dB higher than the mean value of the cage experiment (s.e. 0.35, t 7.0, $P < 0.001$). At 120 kHz the mean value for the model is 0.7 dB lower than the experiment (s.e. 0.52, t 1.4, P 0.2).



obtained by confining known numbers of krill in a nylon mesh cage suspended from a moored raft. Downward-directed echo sounder transducers operating at 38 and 120 kHz were aimed at the cage (0.5 m high and 0.5 m diameter). Up to 1,600 krill were placed in the cage producing densities as high as $16,000 \text{ m}^{-3}$, well within the range reported for naturally occurring swarms³. Additional experiments were carried out with no krill in the cage. The transducers were connected to a Simrad EK400 echo sounder, Simrad QD digital integrator and data logger. The echo sounder transceivers were periodically calibrated and transducer beam patterns measured by the moving sphere method¹⁰. A tungsten carbide sphere, positioned close to the main acoustic axes of both transducers and above the cage, provided a continuous check on system stability.

Experiments were run for 15–64 h and the mean echo level recorded every 6 min. The TS that was determined, first by subtraction of echo energy of cage and krill (Table 1), and second by regression of echo energy against krill number (Table 1), is very much lower than calculated values using equations produced by the Biological Investigation of Marine Antarctic Systems and Stocks (BIOMASS) programme (Table 1). At 120 kHz the values are at least 10 dB higher than our experimental ones. Although results at 38 and 50 kHz are not directly comparable, a difference of 25 to 30 dB indicates a considerable inconsistency between the two.

TABLE 1 Summary results of TS of *Euphausia superba* from cage experiments calculated by the methods of subtraction and regression and compared with TS calculated from BIOMASS⁴

Frequency (kHz)	Method	cv (%)	TS	TS - s.d.	TS + s.d.
120	Subtraction	31	-76.1	-77.7	-74.9
	Regression	33	-78.6	-80.3	-77.4
	BIOMASS		-65.9		
38	Subtraction	47	-85.1	-87.9	-83.4
	Regression	46	-89.4	-92.1	-87.7
50	BIOMASS		-60.4		

In the subtraction method $TS = 10 \log ((I_k - I_c)/N)$ where I_k is the relative echo energy due to the cage and krill, I_c the echo energy due to the cage alone, and N is the number of krill in the cage. This approach assumes that the echo energy due to a group of scatterers is equal to the sum of the individual scatterer contributions to the echo energy. In the regression method the single krill target strength was estimated from the least squares regression of I_k on N for all experiments. Mean length of krill in the experiments ranges from 29.6 mm to 38.9 mm, overall mean is 32.6 mm. cv(%), Coefficient of variation.

The difference between the TS values calculated by subtraction and regression is small relative to the difference between these and the BIOMASS values⁴. It is still nonetheless disconcertingly large. This may be due in part to the relatively large and variable contribution of the empty cage. Subsidiary causes of variability may be due to the degree of aggregation, orientation and size of krill. Video recordings of the krill revealed frequent concentration of the animals into a small fraction of the cage, even at the highest densities used. While some of this behaviour was probably induced by water flow through the cage, the general trend of daytime concentration and dusk dispersal was similar to behaviour observed in the open ocean¹¹. The dominant characteristics affecting TS of marine organisms are gas inclusions, which krill do not possess, and contrasts in density, longitudinal-wave sound speed, and compressibility with respect to sea water. Using the scattering model of Greenlaw⁸ we calculated a theoretical TS and compared it with the subtraction values from the cage experiments (Fig. 1). Although agreement between model and experimental values at 120 kHz is greater than at 38 kHz, the model values are very much closer to our experimental values than to the BIOMASS values (Fig. 1). The model is by necessity oversimplified with respect to recent models¹² in which orientation has a key role, and further study on orientation, degree of aggregation, and biological condition is likely to be most fruitful in connecting measurement and theory.

The BIOMASS equations⁴ show a marked change in TS with krill size. Unfortunately the size spectrum of krill used in our experiments was very small (23–45 mm). Krill may grow to more than 60 mm long³ so our results represent a small part of the overall size spectrum of krill present in the wild. We therefore consider it inadvisable to extrapolate our results to provide TS-to-size relationship over the full size range of krill. Both the implications of these results are clear: values of TS used previously are far too high and standing stock estimates which have relied on them underestimate krill biomass by at least an order of magnitude. We believe that the results reported here, although having some limitations due to the small size range of krill available for the study, provide a far more reliable basis for future estimates of biomass. □

Received 17 October 1989; accepted 13 March 1990.

1. Laws, R. M. *Am. Sci.* **73**, 26–40 (1985).
2. Report of the Seventh Meeting of the Scientific Committee of the Commission for the Conservation of Antarctic Marine Living Resources, SC-CAMLR-VII, XXXX (CCAMLR, Hobart, 1988).
3. Miller, D. G. M. & Hampton, I. *BIOMASS Sci. Ser.* **9**, (1989).
4. Report on Post-FIBEX Acoustic Workshop, Frankfurt, Federal Republic of Germany, *BIOMASS Report Series No. 40* (1986).
5. Croxall, J. P., Prince, P. A. & Ricketts, C. in *Antarctic Nutrient Cycles and Food Webs* (ed. Siegfried, W. R., Condy, P. R. & Laws, R. M.) 516–533 (Springer, Berlin, 1985).

6. Everson, I. & Bone, D. G. *Polar Biol.* **6**, 83–90 (1986).
7. Foote, K. G. *J. Acoust. Soc. Am.* **73**, 1932–1940 (1983).
8. Greenlaw, C. F. *Limnol. Oceanogr.* **24**, 226–242 (1979).
9. Foote, K. G., Everson, I., Watkins, J. L. & Bone, D. G. *J. Acoust. Soc. Am.* **87**, 16–24 (1990).
10. Foote, K. G., Knudsen, H. P., Vestnes, G., MacLennan, D. N. & Simmonds, E. J. *Calibration of Acoustic Instruments for Fish Density Estimation: A Practical Guide* (ICES, Copenhagen, 1987).
11. Everson, I. *J. Plankton. Res.* **4**(1), 155–162 (1982).
12. Stanton, T. K. *J. Acoust. Soc. Am.* **83**, 55–63 (1988).
13. Kogeler, J. W., Falk-Petersen, S., Kristensen, A., Pettersen, F. & Dalen, J. *Polar Biol.* **7**, 231–235 (1987).

Gliding behaviour and palaeoecology of the alleged primate family Paromomyidae (Mammalia, Dermoptera)

K. Christopher Beard

Section of Vertebrate Fossils, Carnegie Museum of Natural History,
4400 Forbes Avenue, Pittsburgh, Pennsylvania 15213, USA

THE extinct Paromomyidae, known from early Tertiary strata of North America^{1–6} and Europe^{7–8}, have traditionally been viewed as an archaic radiation of the order Primates^{1–11}. Hypotheses about the phylogenetic position and palaeobiology of these animals have, however, been highly constrained because postcranial fossils of paromomyids have not been positively identified until now. Newly discovered fossils representing the paromomyid genera *Phenacolemur* and *Ignacius* show that these animals share functionally important postcranial synapomorphies with extant *Cynocephalus* (the flying lemur), the only living member of the mammalian order Dermoptera. These traits strongly suggest that paromomyids possessed a patagium, or gliding membrane, which was homologous to that of *Cynocephalus*. Paromomyids therefore provide both the earliest evidence for the evolution of gliding behaviour in mammals and the only fossil record currently recognized for the eutherian order Dermoptera¹².

Two dentally associated partial skeletons of the paromomyid genus *Phenacolemur* have recently been recovered from early Eocene (Wasatchian) rocks in the Bighorn and Clark's Fork

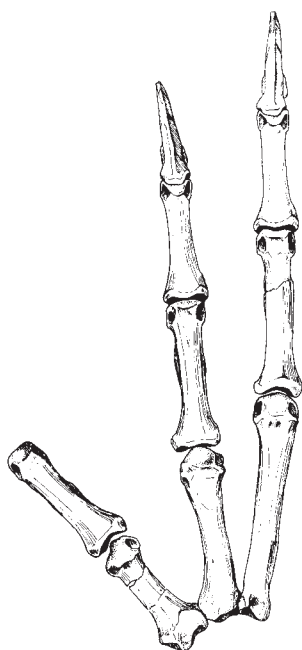


FIG. 1 Right partial hand skeleton of the plesiadapid *Nannodectes intermedius* (USNM 442229). Note relative proportions of phalangeal elements.



FIG. 2 Proximal (bottom; USNM 442248) and intermediate (top; USNM 442254) manual phalanges of the paromomyid *Phenacolemur simonsi*.

Basins of northwestern Wyoming. These specimens, UM 66440 and UM 86352 (which represent a single individual of *P. praecox*) and USGS 17847, have allowed the identification of additional, dentally unassociated postcranial fossils representing both *Phenacolemur* and the closely related paromomyid *Ignacius* from an early Wasatchian faunule in the Clark's Fork Basin^{13,14}.

Perhaps the most unusual aspect of the postcranial anatomy of paromomyids revealed by the new specimens is the osteology of the manus and pes. One way in which paromomyids differ from plesiadapids (believed to be closely related to paromomyids on the basis of craniodental morphology^{4,9–10}) is in possessing intermediate phalanges that are much longer than their corresponding proximal phalanges (Figs 1–3). This very uncommon condition is found among living mammals only in the dermopteran *Cynocephalus* and in edentates (sloths, armadillos and anteaters) (Fig. 3). The overall anatomy of the phalanges of edentates is, however, very different from that of *Cynocephalus* and paromomyids, and their unusual phalangeal proportions seem to result from their peculiarly stub-like proximal phalanges, rather than from elongation of the intermediate phalanges *per se* (Fig. 3). By contrast, phalangeal morphology is extremely similar in paromomyids and extant *Cynocephalus*.

Given this close anatomical similarity, the functional significance of the greatly elongated intermediate phalanges of paromomyids can be assessed in light of the anatomy of extant *Cynocephalus*. In *Cynocephalus* the patagium extends between the individual digits of both the manus and the pes, so that pronounced interdigital webbing of the patagium occurs from the level of the metacarpo- or metatarsophalangeal joints (proximally) to the level of the distal interphalangeal joints (distally)¹⁵. Although the capacity for gliding has evolved independently at least six times in mammals, only in *Cynocephalus* does the patagium form this interdigital webbing¹⁶; for this reason *Cynocephalus* has been described as a 'mitten glider'¹⁶.

Because of the interdigital webbing of the patagium in *Cynocephalus*, the degree of flexion-extension of the proximal interphalangeal joints, especially of the manus, should have a direct effect on the amount of billowing to which the patagium is subjected. The great elongation of the intermediate phalanges in *Cynocephalus* amplifies this relationship, and therefore seems to be functionally related to enhancing voluntary control over the degree of patagial billowing. The latter factor greatly influences the aerodynamics of gliding in mammals¹⁷.

The new evidence presented here strongly indicates that paromomyids also possessed a patagium with significant interdigital webbing, as occurs only in *Cynocephalus* among extant mammalian gliders. As noted previously, plesiadapids lack the specializations in phalangeal anatomy shared by paromomyids and *Cynocephalus*. Therefore, there is no evidence that plesiadapids possessed a patagium. In a phylogenetic context, the evolution of a patagium for gliding locomotion and the osteological features that are functionally associated with this development imply that paromomyids comprise the sister group of *Cynocephalus*, thus forming a clade that excludes the family Plesiadapidae (Fig. 4). Further postcranial synapomorphies also link paromomyids with extant *Cynocephalus*¹³. These traits include a major reorganization of the proximal carpus, in which the lunate is positioned distal to the scaphoid rather than ulnar to it (thus allowing the triquetrum to articulate with both lunate and scaphoid on its radial aspect), and strong development of a secondary synovial articulation between the posterior side of the sustentaculum tali and the plantar aspect of the medial astragalar pillar. Although these and other postcranial characters¹³ provide substantial additional support for the phylogenetic relationships depicted in Fig. 4, they are not discussed further here because they have no direct functional relationship with gliding behaviour.

In addition to forming the basis for this novel phylogenetic hypothesis, the new postcranial fossils of paromomyids provide fresh insight into their palaeobiology. For example, it has been known for some time that the dentition of paromomyids is strongly convergent on that of the extant sugar-glider *Petaurus*^{18–20}, a phalangeroid marsupial that feeds extensively on tree exudates^{21,22}. The fossil evidence reported here suggests that paromomyids were a much closer ecological analogue to extant *Petaurus* than had previously been imagined, as they apparently evolved a patagium convergent on that of extant sugar-gliders in addition to the dental features noted previously.

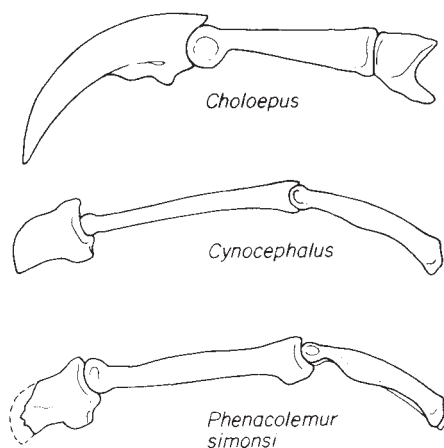


FIG. 3 Manual phalanges of the edentate *Choloepus* (two-toed sloth), the dermopteran *Cynocephalus* (flying lemur), and the paromomyid *Phenacolemur* in lateral view (not to same scale). Compare with Fig. 1.

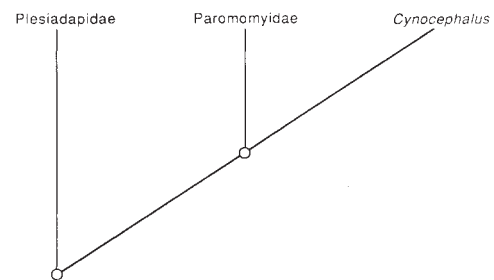


FIG. 4 Cladogram illustrating phylogenetic position of Paromomyidae, based on characters discussed in the text. Note that the possible sister-group relationship between Plesiadapidae and a clade comprising Paromomyidae and *Cynocephalus* has not been established here (but see ref. 13); rather, plesiadapids are simply used as an outgroup for the latter clade here.

The preceding hypothesis is corroborated by the pattern of individual abundance and stratigraphic longevity shown by paromomyids. Unlike plesiadapids, which often formed a dominant component of Palaeocene mammalian assemblages in both Europe²³ and North America^{24,25}, paromomyids were never very abundant, as might be expected of animals with a highly specialized ecological niche. Likewise, while plesiadapids became extinct near the Palaeocene–Eocene boundary in North America^{6,26}, paromomyids persisted on that continent into the Duchesnean Land-Mammal Age (late Eocene)^{5,27}. It has been suggested²⁶ that the extinction of plesiadapids in North America may have resulted from competitive exclusion by early rodents, which immigrated into North America from Asia at the beginning of Clarkforkian time (late Palaeocene)²⁸. The ecological specializations of paromomyids discussed here evidently resulted in minimal competition between paromomyids and early rodents, allowing the two taxa to coexist in North America for some 18 million years²⁹. The eventual extinction of paromomyids in North America occurred near the Eocene–Oligocene boundary, at which time rapid climatic and floral changes³⁰ probably resulted in the deterioration of habitats required by such obligatory forest-dwelling mammals. □

Received 18 January; accepted 6 March 1990.

- Matthew, W. D. *Bull. Am. Mus. nat. Hist.* **34**, 429–483 (1915).
- Gidley, J. W. *Proc. US natn. Mus.* **63**, 1–38 (1923).
- Simpson, G. G. *Bull. Am. Mus. nat. Hist.* **105**, 411–442 (1955).
- Bown, T. M. & Rose, K. D. *Folia primatol.* **26**, 109–138 (1976).
- Krishtalka, L. *Ann. Carnegie Mus.* **47**, 335–360 (1978).
- Rose, K. D. & Bown, T. M. *J. vert. Paleont.* **2**, 63–69 (1982).
- Russell, D. E., Louis, P. & Savage, D. E. *Univ. Calif. Publ. geol. Sci.* **73**, 1–46 (1967).
- Godinot, M. *Folia primatol.* **43**, 84–96 (1984).
- Szalay, F. S. & Delson, E. *Evolutionary History of the Primates* (Academic, New York, 1979).
- Szalay, F. S., Rosenberger, A. L. & Dagosto, M. *Yb. phys. Anthropol.* **30**, 75–105 (1987).
- Szalay, F. S. *Am. J. phys. Anthropol.* **36**, 59–76 (1972).
- MacPhee, R. D. E., Cartmill, M. & Rose, K. D. *J. vert. Paleont.* **9**, 329–349 (1989).
- Beard, K. C. thesis, The Johns Hopkins University, Baltimore, Maryland (1989).
- Beard, K. C. & Houde, P. *J. vert. Paleont.* **9**, 388–399 (1989).
- Poock, R. I. *Proc. zool. Soc. Lond.* 429–444 (1926).
- Therington, R. W., Jr *Science* **225**, 1048–1050 (1984).
- Therington, R. W., Jr & Heaney, L. R. *J. Mammal.* **62**, 101–114 (1981).
- Gingerich, P. D. *J. dent. Res.* **53**, 497 (1974).
- Cartmill, M. *Science* **184**, 436–443 (1974).
- Kay, R. F. & Cartmill, M. *J. hum. Evol.* **6**, 19–53 (1977).
- Wakefield, N. A. *Vict. Nat.* **87**, 221–236 (1970).
- Smith, A. P. *J. Anim. Ecol.* **51**, 149–166 (1982).
- Russell, D. E. *Mém. Mus. natn. Hist. nat. (Sér. C)* **13**, 1–324 (1964).
- Gingerich, P. D. *Univ. Mich. Pap. Paleont.* **15**, 1–141 (1976).
- Rose, K. D. *Univ. Mich. Pap. Paleont.* **26**, 1–197 (1981).
- Maas, M. C., Krause, D. W. & Strait, S. G. *Paleobiology* **14**, 410–431 (1988).
- Storer, J. E. *Dakoterra* **3**, 108–117 (1987).
- Krause, D. W. *Univ. Wyoming Contr. Geol. spec. Pap.* **3**, 95–117 (1986).
- Berggren, W. A., Kent, D. V., Flynn, J. J. & Van Couvering, J. A. *Geol. Soc. Am. Bull.* **96**, 1407–1418 (1985).
- Wing, S. L. *Ann. Missouri Bot. Gard.* **74**, 748–784 (1987).

ACKNOWLEDGEMENTS. I am grateful to Drs T. M. Bown, P. D. Gingerich, K. D. Rose, R. W. Therington, Jr, A. Walker, and especially P. Houde for access to fossil and recent material, and Drs R. F. Kay, L. Krishtalka, and K. D. Rose for critical review of the manuscript. Figures 1–3 were prepared by Ms E. Kasmer. This work was supported by the NSF.

Eocene plesiadapiform shows affinities with flying lemurs not primates

Richard F. Kay*, Richard W. Thorington Jr† & Peter Houde‡

* Department of Biological Anthropology and Anatomy, Duke University Medical Center, Durham, North Carolina 27710, USA

† Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA

‡ Department of Biological Sciences, Princeton University, Princeton, New Jersey 08544, USA

PLESIADAPIFORMES, of the North American and European Paleogene, is often identified as a sister group of primates. This hypothesis is based on several proposed anatomical synapomorphies linking the best-known plesiadapiform families, Plesiadapidae, and Paromomyidae with Eocene primates^{1–5}. The first well-preserved skull of *Ignacius graybullianus*, an early Eocene paromomyid plesiadapiform, clarifies and corrects previous cranial reconstruction based on more fragmentary material^{3,6,7}. The new material indicates Plesiadapiformes are not Primates. Rather, several synapomorphies argue for a closer phylogenetic relationship between Plesiadapiformes and *Cynocephalus*, the extant flying lemur (order Dermoptera). In view of the finding that “archaic” primates are not cladistic Primates, the recently coined taxon “Euprimates” should be discarded. No support is lent by cranial anatomy to the hypothesis that Primates, tree shrews, bats and dermopterans form a clade Archonta.

The new specimen of *Ignacius graybullianus*, USNM 421608, is a nearly complete skull collected by P. Houde from early Wasatchian (Wa₁, earliest Eocene⁸) rocks of the Willwood Formation, Clark's Fork Basin, Park County, Wyoming. The USNM 421608 skull (Figs 1 and 2) is slightly crushed interorbitally and the face is slightly displaced dorsally relative to the braincase. The premaxillae were lost during deposition, but the rest of the skull, including the basicranium, is complete. Here we focus on the basicranial anatomy considered crucial for phylogenetic inference. Several plesiadapiform basicrania have been reported previously. These include poorly preserved material of the paromomyids *Ignacius*⁶ and *Phenacolemus*³, and several better preserved specimens of the plesiadapid *Plesiadapis*^{1,2}.

The large auditory bulla is completely ossified (Figs 1–3). It is greatly expanded medially and balloons caudally over the basioccipital. The bulla is composed of the entotympanic, not of the basisphenoid, basioccipital, ectotympanic, or, as in primates, the petrosal^{5,9,10}. This is evidenced by the following observations: (1) sutures separate the medial, rostral and caudal bullar walls from the basisphenoid and basioccipital; (2) laterally, sutures are visible between the collar-shaped ectotympanic and the bulla; (3) inside the right bulla on the roof of the tympanic cavity, the promontorial part of the petrosal is separated from the bulla medially by a suture. That this is a suture and not a fracture or other artefact is demonstrated by structure and symmetry. This suture is formed by an overlapping of the petrosal element onto the bone of the bulla. The bones thin out at their junction rather than butting bluntly against one another as might appear in postmortem breakage. Furthermore, the position of the suture is identical on both sides of the skull as viewed directly ventrally on the right side (Fig. 3) and through a rostromedial break in the left bulla. The bulla of *Plesiadapis* may have formed in a similar manner followed by full co-ossification of the entotympanic–petrosal suture¹¹.

Having an auditory bulla formed wholly or in part by an entotympanic is a key distinction between Plesiadapiformes and Primates, which have a petrosal bulla^{5,9}. If the bulla was mem-

branous in primitive eutherians^{12,13}, the presence of an entotympanic bulla in *Ignacius* would be a derived character that rules *Ignacius* and other Plesiadapiformes out of the ancestry of Primates. Bullar entotympanics have evolved in many mammalian groups¹⁰. However, an unusual similarity in the entotympanics of *Ignacius* and the living dermopteran *Cynocephalus* is that the entotympanic contacts the basioccipital ventromedially¹⁴. This is a distinction from the entotympanics of tree shrews, for example, in which the only bony contact medially is with the petrosal.

Previous proposals of a significant intracranial blood supply for paromomyids by way of either the internal carotid artery^{3,9} or the ascending pharyngeal artery⁶ are ruled out by the following observations. (1) The carotid foramen is too small to have conducted an internal carotid artery of significant size. The diameter of the carotid foramen in *Ignacius* is comparable to that of lorises, in which the internal carotid artery involutes during development and only the internal carotid nerve passes through^{15,16}. (2) There are no grooves or tubes for the promontory division of the internal carotid artery on the promontorium. It has been suggested that in *Phenacolemus* the promontorial branch travelled through a longitudinal septum running anteriorly from the promontorium³. In *Phenacolemus*, however, this strut lacks a lumen for transmitting an artery¹¹. In USNM 421608 the homologous strut is broken rostrally and the broken part likewise shows no lumen. Moreover, the longitudinal septum projects too far laterally for it to transmit an artery to the cerebral arterial circle. (3) There are no grooves on the basioccipital or basisphenoid along the cranial base, nor is there a middle lacerate foramen to transmit an ascending pharyngeal artery from the neck to the inside of the braincase as in cheirogaleids and lorises¹⁵. Foramina and grooves for these structures previously identified on another more fragmentary

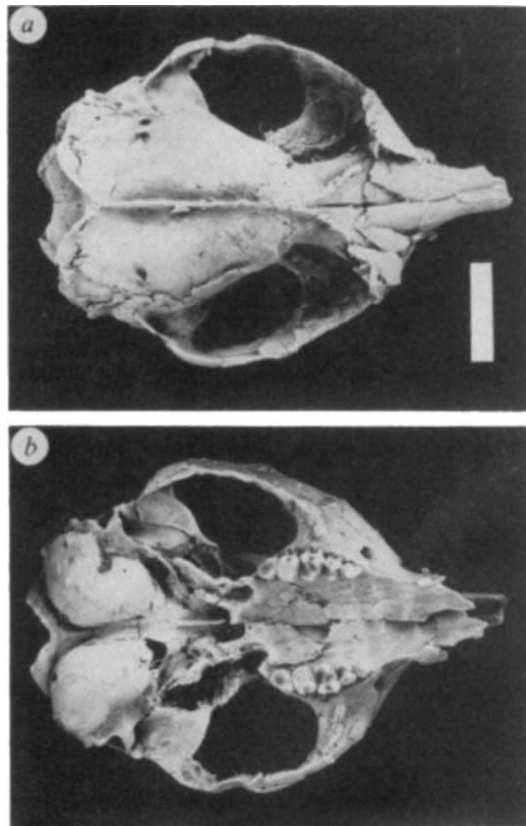


FIG. 1 Photographs of *Ignacius graybullianus*, USNM 421608, from Houde site 24 (near UM (University of Michigan) site 125 (ref. 8)), SW 1/4 of NW 1/4 of section 24, T56N R102W, Clark Quadrangle N4445-W10900/15 1950. a, Dorsal view; b, ventral view. Scale bar, 10 mm. Photography, V. E. Krantz.

skull of *Ignacius*⁶ are perhaps for the inferior petrosal venous sinus that ran inside the braincase or are artefacts of the separation of bones on the cranial base during fossilization⁵. By the process of elimination, the principal intracranial blood supply of *Ignacius* was probably from the vertebral arteries by way of the foramen magnum. This means of supplying intracranial blood, together with the exclusion of an internal carotid artery and of an ascending pharyngeal artery from such a role, is a synapomorphy of *Plesiadapis*¹¹, *Ignacius* and extant Dermoptera^{14,17}.

The anatomical features described above provide new insight into the phylogenetic position of Plesiadapiformes. Plesiadapiformes lack convincing craniodental or postcranial

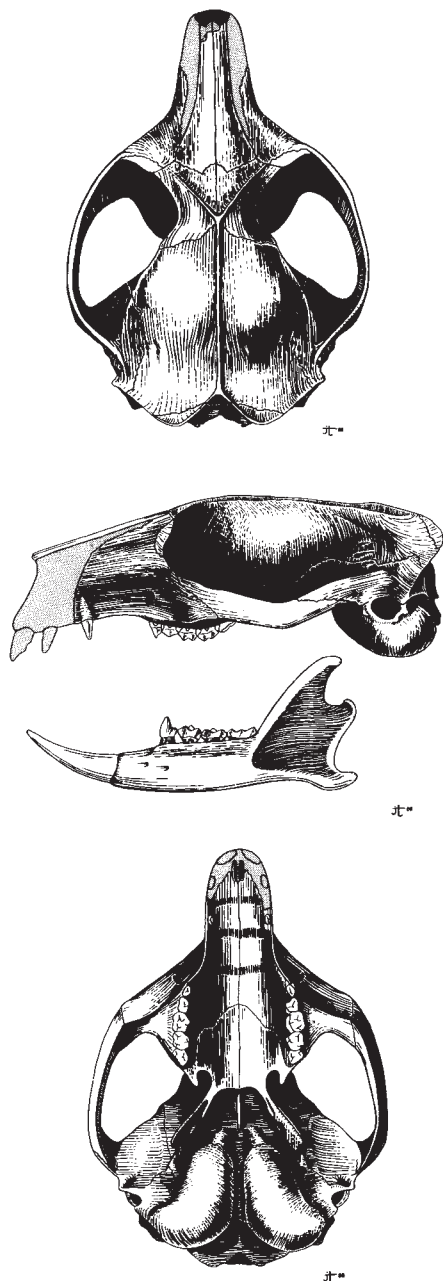


FIG. 2 Reconstruction of *Ignacius* based principally on USNM 421608 and the premaxilla (represented by grey stipple) of a previously described specimen of this species⁷. The lower jaw in the reconstruction is based on an immature specimen from Houde site 14 (probably within UM site SC 26) also from early Wasatchian rocks of the Willwood Formation, Clark's Fork Basin, Wyoming. Drawing, J. Trecartin.

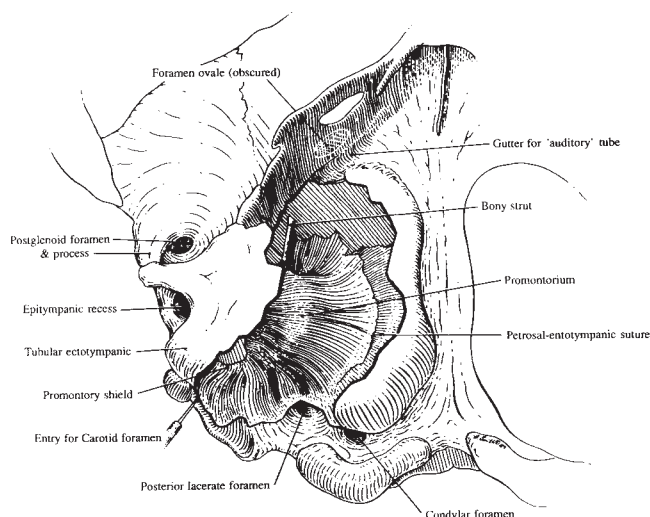


FIG. 3 The right auditory region of USNM 421608 from a ventral perspective. Part of the auditory bulla has been removed. Drawing, B. Smith.

synapomorphies with Primates. They do not have such distinctively primate features as a postorbital bar, orbital convergence with narrowed interorbital space, an enlarged brain, a large internal carotid artery in a promontorial position enclosed in a bony canal, or a petrosal auditory bulla. Previously suggested postcranial synapomorphies^{5,9} linking Plesiadapiformes with Primates also seem to be unfounded¹⁸⁻²⁰. About the only derived similarity between Plesiadapiformes and some early Eocene Primates is the presence of a *Nannopithecus*-fold on the upper molars²¹. However, *Nannopithecus*-folds are poorly developed in the earliest primates of modern aspect (such as *Cantius* or *Teilhardina*) suggesting that the feature evolved independently in Plesiadapiformes and Primates. In the absence of any convincing evidence linking Plesiadapiformes with primates of modern aspect, there is now no need for the term Euprimates (see ref. 5) as distinct from Primates.

The hypothesis that Plesiadapiformes may form a monophyletic clade with Dermoptera^{19,20} is supported by the following cranial synapomorphies. (Recent evidence shows that Eocene Plagiomenidae have no close relationship to Dermoptera as previously supposed²².) Extant Dermoptera and Plesiadapiformes share the collar-shaped ectotympanic, a partially involuted internal carotid arterial system, a bulla composed partly of the entotympanic, and an entotympanic that contacts the basioccipital medially. The last two features are demonstrated in Plesiadapiformes only for *Ignacius*, but the reported absence of similar sutural patterns in the fully ossified bulla of *Plesiadapis* does not rule out the possibility of an entotympanic in this taxon as well^{6,11}. Further, we find no features of the cranial anatomy to support the hypothesis that Primates, tree shrews, bats and dermopterans form a clade Archonta. A hypothetical ancestor of such a group would have resembled primitive eutherians in having an unossified cranial bulla and unreduced internal carotid system. In summary, the new cranial and postcranial fossils of paromomyid Plesiadapiformes discussed here and elsewhere^{19,20} indicate that these animals were not 'archaic primates' as commonly believed^{1-5,9,21}. Rather, they appear to be more aptly termed 'archaic dermopterans'. □

Received 18 January; accepted 6 March 1990.

1. Gingerich, P. D. *Univ. Mich. Pap. Paleont.* **15**, 1-141 (1976).
2. Russell, D. E. *Mem. Mus. nat. Hist.* **13**, 1-324 (1964).
3. Szalay, F. S. *Am. J. phys. Anthropol.* **36**, 59-76 (1972).
4. Szalay, F. S. *Bull. Am. Mus. nat. Hist.* **140**, 193-330 (1969).
5. Szalay, F. S., Rosenberger, A. L. & Dagosto, M. *Yb. phys. Anthropol.* **30**, 75-105 (1987).
6. MacPhee, R. D. E., Cartmill, M. & Gingerich, P. D. *Nature* **301**, 509-511 (1983).
7. Rose, K. D. & Gingerich, P. D. *Contrib. Mus. Paleont. Univ. Mich.* **24**, 181-189 (1976).

8. Gingerich, P. D. *Wyoming Geol. Assn. Guidebk* **34**, 185-195 (1983).
9. Szalay, F. S. & Delson, E. *Evolutionary History of the Primates* 1-580 (Academic, New York, 1979).
10. Novacek, M. J. *Mammal. Rev.* **7**, 131-149 (1977).
11. MacPhee, R. D. E. & Cartmill, M. in *Comparative Primate Biology: Systematics, Evolution and Anatomy* (eds Swindler, D. & Erwin, J.) 219-275 (Alan R. Liss, New York, 1986).
12. Novacek, M. J. *Bull. Am. Mus. nat. Hist.* **183**, 1-112 (1986).
13. Szalay, F. S. *J. hum. Evol.* **6**, 3-18 (1977).
14. Hunt, R. M. & Korth, W. W. *J. Morph.* **164**, 167-211 (1980).
15. Cartmill, M. in *Phylogeny of the Primates: A Multidisciplinary Approach* (eds Luckett, W. P. & Szalay, F. S.) 313-354 (Plenum, New York, 1975).
16. Conroy, G. C. & Packer, D. J. *Folia primatol.* **35**, 237-247 (1981).
17. Wible, J. R. *Zool. J. Linn. Soc.* **91**, 107-135 (1987).
18. Wible, J. & Covert, H. H. *J. hum. Evol.* **16**, 1-22 (1987).
19. Beard, K. C. thesis, The Johns Hopkins University, Baltimore, MD (1989).
20. Beard, K. C. *Nature* **345**, 340-341 (1990).
21. Hoffstetter, R. C. *r. heb. Seanc. Acad. Sci., Paris (Ser. II)* **302**, 43-45 (1986).
22. MacPhee, R. D. E., Cartmill, M. & Rose, K. D. *J. vert. Paleont.* **9**, 329-349 (1989).

ACKNOWLEDGEMENTS. This specimen was collected with the assistance of Anne Houde. The work was supported by a Scholarly Studies Grant of the Smithsonian Institution to Dr Storrs Olson. R.F.K. was supported by a Smithsonian Institution fellowship.

Limited heterogeneity of rearranged T-cell receptor V α transcripts in brains of multiple sclerosis patients

Jorge R. Oksenberg*, Simon Stuart†, Ann B. Begovich‡, Robert B. Bell*, Henry A. Erlich‡, Lawrence Steinman*§|| & Claude C. A. Bernard†

Departments of *Neurology, §Genetics and Pediatrics, Stanford University, Stanford, California 94305, USA

†Neuroimmunology Laboratory, LaTrobe University, Victoria 3083, Australia

‡Department of Human Genetics, Cetus Corporation, Emeryville, California 94608, USA

THE identification of activated T cells in the brain of individuals with multiple sclerosis (MS) indicates that these cells are critical in the pathogenesis of this disease. In an attempt to elucidate the nature of the lymphocytic infiltration, we used the polymerase chain reaction to amplify T-cell antigen receptor (TCR) V α sequences from transcripts derived from MS brain lesions. In each of three MS brains, only two to four rearranged TCR V α transcripts were detected. No V α transcripts could be found in control brains. Sequence analysis of transcripts encoded by the V α 12.1 region showed rearrangements to a limited number of J α region segments. These results imply that TCR V α gene expression in MS brain lesions is restricted.

Multiple sclerosis is an inflammatory disease of the central nervous system (CNS), characterized by myelin destruction¹⁻³. In the brain, there is an accumulation of macrophages, plasma cells, major histocompatibility complex (MHC) class II positive antigen-presenting cells and activated cytokine-secreting T cells⁴⁻⁸. Several lines of evidence indicate that T lymphocytes migrate from the peripheral blood to the CNS compartment and participate directly in the formation of brain lesions⁹⁻¹¹. There is also evidence of oligoclonality in T lymphocytes within the cerebrospinal fluid of MS patients¹². In addition, TCR V α and V β genes have been shown to contribute to the genetic control of susceptibility to this disease¹³⁻¹⁵.

To examine the expression of TCR genes at the site of disease, messenger RNA isolated from demyelinating brain plaques from three MS patients with chronic progressive disease, and from three control brains (non-MS) was used to synthesize complementary DNA. These cDNAs were then subjected to enzymatic gene amplification by the polymerase chain reaction (PCR) method^{16,17} using specific TCR primers. The results of such an amplification using primers for the TCR V α 12.1 family

TABLE 1 T-cell receptor α -primers

Primer	Clone	Sequence	Family members
V α 1	HAP 10	5'-CTGAGGTGCAACTACTCA-3'	1.1, 1.2, 1.3
V α 2	HAP 26	5'-GTGTTCCAGAGGGAGCCATTGCC-3'	2.1, 2.2
V α 3	HAP 05	5'-GGTGAACAGTCAACAGGGAGA-3'	3.1
V α 4	HAP 08	5'-ACAAGCATTACTGTACTCTCA-3'	4.1
V α 5	HAP 35	5'-GGCCCTGAACATTCAGGA-3'	5.1
V α 6	HAP 01	5'-GTCACCTTCTAGCCTGCTGA-3'	6.1
V α 7	HAP 21	5'-AGGAGCCATTGTCCAGATAAA-3'	7.1, 7.2
V α 8	HAP 41	5'-GGAGAGAATGTGGAGCAGCATC-3'	8.1, 8.2
V α 9	HAP 36	5'-ATCTCAGTCTGTGATAATA-3'	9.1
V α 10	HAP 58	5'-ACCCAGCTGGTGGAGCAGAGCCCT-3'	10.1
V α 11	HAP 02	5'-AGAAAGCAAGGACCAAGTGT-3'	11.1
V α 12	PGA 5	5'-CAGAGGTAACCTCAAGCGCAGACT-3'	12.1
V α 13	AB 11	5'-GCTTATGAGAACACTGCGT-3'	13.1
V α 14	AB 21	5'-GCAGCTTCCCTCCAGCAAT-3'	14.1
V α 15	AC 24	5'-AGAACCTGACTGCCAGGAA-3'	15.1
V α 16	AE 212	5'-CATCTCCATGACTCATATGA-3'	16.1
V α 17	AF 211	5'-GACTATACTAACAGCATGT-3'	17.1
V α 18	AC 9	5'-GTCTCAGCAATGACAAAG-3'	18.1
C α	PGA 5	5'-AATAGTCGAGACACTGTCTGGA-3'	C α

The size of amplified products using 5' V α and 3' C α primers ranged from about 320 to 410 base pairs.

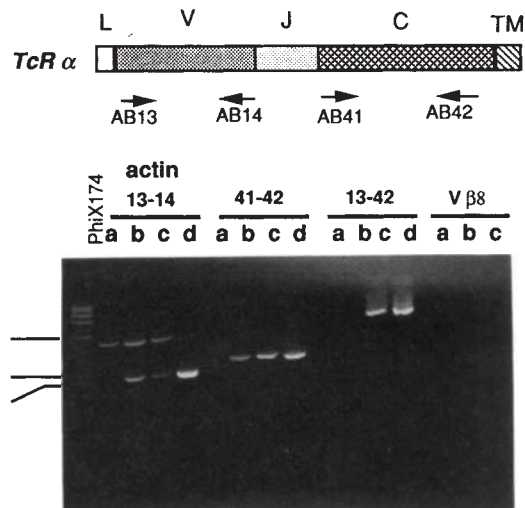


FIG. 1 Brain TCR amplification of MS patient 1. Lanes a: control brain cDNA; b, MS parietal region brain cDNA; c, MS occipital region brain cDNA; d, PGA5, a full-length TCR α cDNA¹⁸. Complementary DNA (2 μ l) was combined in a 100- μ l reaction volume, with 1 unit of DNA *Taq* polymerase (Perkin Elmer-Cetus), 10 μ l 10 $\times$ reaction buffer, 50 μ M each of deoxynucleoside triphosphates, and 1 μ M of each primer. The PCR profile used was: denaturation 95 $^{\circ}$ C for 60 s, annealing 45 $^{\circ}$ C for 60 s and extension 72 $^{\circ}$ C for 60 s, for 35 cycles on a DNA Thermal Cycler (Perkin Elmer-Cetus). One tenth of each sample was independently run in a 4% Nusieve agarose gel (Fmc) and an appropriate size fraction was excised from the gel. The agarose piece was frozen and thawed 3 times, and 2 μ l supernatant directly reamplified with the same primers for an additional 25 cycles. A 500-bp actin fragment was successfully amplified from brain cDNA (lanes a-c), but not from the PGA5 control (lane d) using the following primers: 5'-ACGAAGACGACACC-GCCCTCG-3', 5'-CACGTTGTGGGTGACGCCGTC-3'. V α and C α transcripts were amplified from both MS brain cDNA and PGA 5 templates, but not from the control MS brain cDNA with primers AB 13-14 (5'-CAGAAGGTAACCTG-CAGCGCAGACT-3', 5'-TTGGGGATCCAGAGCACAGAAGTATACTGC-3'), which includes *Pst*I and *Bam*HI restriction sites and define a 286-bp fragment of the V α 12.1 region gene; and AB 41-42 (5'-CAGAACCCTGACCCCTGCCG-TGTAC-3', 5'-GTGTCACAGTTAGGTTGCTATCTGT-3', which includes a *Sal*I site and defines a 340-bp fragment of the C α region transcript) respectively. Note that rearranged TCR α sequences could be amplified from cDNA of the MS brain prepared from the occipital region (lane c) using the V α 12.1 primer AB 13 and C α primer AB 42.

|| To whom correspondence should be addressed.

TABLE 2 T-cell receptor V α expression in brain plaques of multiple sclerosis patients

	V α 1	V α 2	V α 3	V α 4	V α 5	V α 6	V α 7	V α 8	V α 9	V α 10	V α 11	V α 12	V α 13	V α 14	V α 15	V α 16	V α 17	V α 18	V β 8	Actin
Experiment no. 1																				
MS Br1	383	4,640	760	520	240	850	826	1,566	450	45,860	5,430	36,380	3,618	367	280	289	226	442	170	104,450
MS Br2	140	824	523	310	830	415	660	23,200	1,750	29,630	623	49,125	456	220	317	12,460	3,572	338	280	79,120
MS Br3	638	313	276	410	817	1,520	210	15,860	16,310	21,200	838	2,050	302	225	462	3,633	482	470	630	58,358
C Br4	235	1,100	135	115	286	7,300	427	960	1,036	317	560	726	485	278	466	630	545	830	900	65,996
C Br5	580	875	180	490	110	846	160	324	780	120	344	138	762	755	876	860	715	570	860	66,393
C Br6	137	290	133	530	836	640	910	110	140	350	670	1,030	1,095	2,000	437	775	240	330	710	139,337
Experiment no. 2																				
MS Br1	1,650	3,956	1,450	790	547	545	1,170	343	1,856	32,870	513	12,978	866	868	3,190	280	1,048	1,127	440	38,593
MS Br2	967	340	1,419	1,575	3,866	2,837	1,848	13,373	2,974	17,337	1,550	33,020	1,487	1,072	3,148	17,968	1,446	980	1,338	32,460
MS Br3	666	726	1,198	790	1,769	258	576	35,270	18,990	19,138	948	2,690	587	880	815	945	946	1,570	630	22,415
C Br4	1,507	660	1,740	1,790	553	706	4,540	4,410	1,333	584	919	765	860	206	590	713	2,748	526	864	31,285
C Br5	896	1,670	2,370	5,000	2,826	418	862	8,175	2,048	1,307	1,734	836	737	1,040	2,097	2,925	1,025	5,276	4,478	33,018
C Br6	883	1,727	716	865	610	1,334	9,514	1,033	1,256	1,130	636	170	4,636	1,300	1,930	1,167	764	5,915	370	29,451
PBL(PHA)	9,434	19,464	8,288	18,434	18,820	10,483	12,800	14,886	13,980	23,040	11,448	16,968	16,536	17,750	30,512	16,544	21,132	19,732	ND	ND

Samples were taken from brain plaques of three patients with chronic progressive MS, and three controls (non-MS). Total RNA and cDNA (from 5 μ g RNA) were prepared according to standard procedures¹³. Control cDNA was also prepared from 1 μ g RNA isolated from a pool of peripheral blood lymphocytes from five different individuals, stimulated with 3 μ g ml⁻¹ of PHA. cDNAs were amplified with TCR V α -C α or actin primers for 40 cycles in the presence of 10 μ Ci of α -[³²P]dATP (Amersham). Samples were analysed by gel electrophoresis with ethidium bromide to identify the specific fragment band (italic figs when a band was clearly visible). After electrophoretic separation, bands were excised and incorporation of radiolabel was determined. Where TCR rearranged bands were absent, an agarose fragment 200–600 bp was excised and counted. An actin band was visualized in all the amplified brain cDNA. Results are expressed in counts per min. All TCR 5' primers amplify TCR sequences from germ line DNA using a specific 3' V α primer for each family. We have detected V-C α rearrangements of all TCR V gene members in a variety of activated T cells including single rearrangements of specific V α members in T-cell clones reactive to pertussis toxin, to Borrelia burgdorferi, as well as the Jurkat T cell line, and rearrangements of all V α members in pooled T cells stimulated by PHA (line 7, experiment 2). ND, Not determined.

on cDNA isolated from the parietal and occipital brain regions from one MS patient, and from the occipital brain region of one control (non-MS) individual, are shown in Fig. 1. Actin sequences were coamplified together with the V α 12.1 gene to monitor the integrity of the cDNAs. Actin could be amplified from the brain cDNAs (500 base pair (bp) fragment (actin primer: lanes a–c) but not from PGA5, a full-length cDNA clone that contains the V α 12.1 segment¹⁸ (lane d). A smaller PCR product corresponding to the V α 12.1 gene was detected in the patient but not in the control sample (282-bp fragment; primer AB13-AB14: lanes b, c and d). To ensure that only the V α 12.1 family was amplified, genomic and brain cDNA PCR products were analysed using restriction endonucleases and were consistent with the known restriction map for V α 12.1. When colonies containing cloned V α PCR products from MS brain cDNA were screened with a V α 12.1 region probe, about 20% were positive. DNA from several of these colonies was sequenced, and was found to be identical with the V α 12.1 sequence¹⁸. Thus, the TCR V α 12.1 restriction fragment-length polymorphism recently associated with MS susceptibility¹⁵, must be in a sequence flanking the V α 12.1 gene. Specific regulation of the α -TCR gene by 3' *cis*-acting enhancers, was recently demonstrated¹⁹.

These experiments indicated that PCR could amplify the receptor transcripts from post-mortem brain samples, starting from nanograms of total RNA without the necessity of *in vitro* expansion of T cells. Similarly, C α sequences were amplified from MS brain cDNAs, but not from the control sample (primers AB41-AB42: lanes b and c, Fig. 1). A subsequent amplification was carried out (primers AB13-AB42: lanes a, b, c and d) using a 5' primer complementary to the V α 12.1 and a 3' primer complementary to the C α TCR regions. These primers, which can amplify only rearranged TCR transcripts, amplified a product of ~680 bp from both the positive control PGA5 sample and cDNA from the occipital region of the MS brain (primers AB13-AB42: lanes c and d) but not from control brain cDNA or cDNA from the parietal region of the MS brain (primers AB13-AB42: lanes a and b). The V α and C α amplifications from the MS brain parietal region most probably represent transcripts from unrearranged chromosomes, as has been found in other cDNA libraries from T cell lines²⁰. No PCR product was observed using primers corresponding to the V β 8 family, even though this TCR V β region was recently reported to be associated with susceptibility to MS (ref. 13).

To substantiate that the DNA produced by PCR was an authentic amplified product of rearranged TCR genes, the PCR products were sequenced using the dideoxy chain termination

method²¹ after double screening of colonies with V α and C α probes. Only two different J regions were seen in the 25 clones examined. Both were different from the PGA5 J α sequence, ruling out the possibility of a 'carryover' contamination. Eleven clones contained the J α O family found in clone HAP 41 (ref.

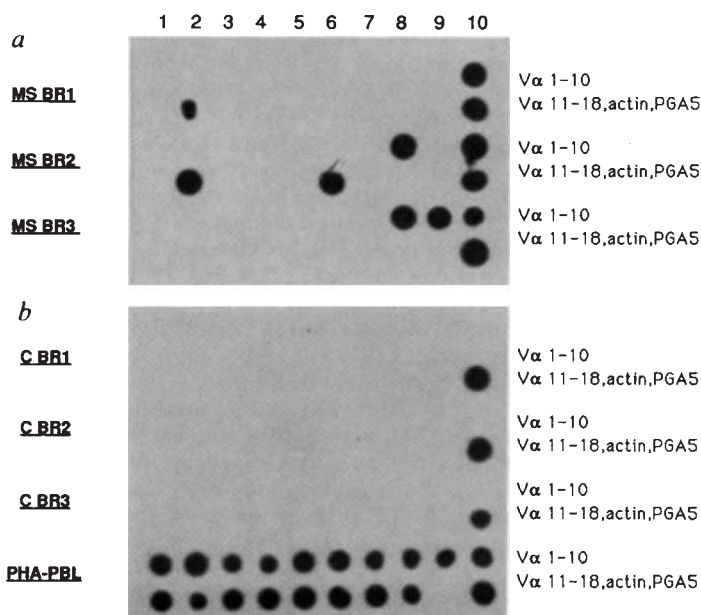


FIG. 2 Detection of PCR-amplified products by dot-blot hybridization using a C α probe. *a*, MS brain (MS BR) amplifications; *b*, control brain (C BR) and PHA-PBL amplifications.

METHODS. The amplified product (10 μ l) was dot-blotted onto transfer membranes (GeneScreen Plus, NEN) after denaturation with 0.4 M NaOH–25 mM Na₂EDTA. Actin PCR products served as negative controls and PGA5 as the positive control. Filters were fixed under ultraviolet light and prehybridized for 3 h at 42 °C in 5 $\times$ SSPE/5 $\times$ Denhardt's solution/salmon sperm DNA (100 μ g ml⁻¹)/0.1% SDS and hybridized overnight at 42 °C with 1 $\times$ 10⁶ c.p.m. per ml of ³²P-kinased probe (5'-AATATCCAGAACCCCTGACCCT-3'). Filters were washed in 1 $\times$ SSPE/0.1% SDS at room temperature, twice for 20 min; and then in 0.1 $\times$ SSPE/0.1% SDS at 45 °C twice for 10 min. Kodak XAR-5 film with Dupont Lightning-Plus intensity screens were used for autoradiography at –70 °C for 30 min. Signals correlate with the presence of a specific band in the ethidium bromide-stained agarose gel and with high [³²P]dATP incorporation during amplification (Table 2).

22) and 14 clones had a previously undescribed J α sequence (GGGTACCGAGATGACGAACCCACCTTTGGGACAGG-CACTCAGCTAAAAGTGGAACTC).

We next asked whether there was a diverse or limited usage of TCR V α gene expression in MS brain lesions. To analyse the use of TCR V α in MS brains, we synthesized 5' PCR primers for the 18 different V α families^{22,23} (Table 1). Optimal conditions for amplification were ascertained using each of the 5' primers (Table 1) in combination with a specific 3' V α primer for each TCR V α family, on genomic DNA (data not shown) and with a common C α primer (AB 51) on reverse-transcribed RNA isolated from phytohaemagglutinin (PHA)-stimulated peripheral blood lymphocytes (Table 2). The results from amplification of MS brain cDNA, using 5' V α primers and the common 3' C α primer AB 51 in the presence of α -[³²P]dATP, showed that in each brain only a few TCR V α gene families are preferentially rearranged and transcribed (Table 2). Visualization of a specific band for actin or rearranged TCR after gel electrophoresis of the amplified product, is indicated by italics in Table 2. In general, high [³²P]dATP incorporation was reproducibly detected in regions of the gel where bands were visualized.

These results were further confirmed by dot blot of the PCR products and hybridization with a radiolabelled oligonucleotide corresponding to a TCR-C α sequence, 5' from the common C α primer (Fig. 2). A clear signal in the dot-blot assay correlates with the presence of a band staining with ethidium bromide in the agarose gel. Products of amplifications with actin primers were used as a negative control; amplification products of clone PGA5 with TCR V α and C α primers were used as a positive control. Expression of TCR V α transcripts is operationally defined by (1) the presence of a rearranged V α -C α band visualized with ethidium bromide, and (2) a signal on dot-blot hybridization of the amplified V α -C α gene with a TCR C α probe, 5' from the C α PCR primer. Thus in MS brain 1, TCR V α 10 and 12 were rearranged; TCR V α 8, 10, 12 and 16 were rearranged in MS brain 2, and V α 8, 9 and 10 were rearranged in MS brain 3. No V α transcripts could be amplified from control brain cDNA, although actin could be amplified. Each of 18 TCR V α rearranged genes were amplified from PHA-stimulated pooled lymphocytes. The V α 10 family was detected in all three MS samples, suggesting that this TCR might be responding to a major epitope of an antigen involved in the pathogenesis of MS. Further experiments will be needed to confirm this hypothesis.

Our results may prove to have therapeutic implications. Studies on TCR gene expression in experimental allergic encephalomyelitis (EAE), a prototypic animal model for autoimmunity induced by T cells, have shown the predominant use of certain TCR α and β products in the immune response to myelin basic protein²⁴⁻²⁷. Treatment of mice *in vivo* who have EAE with monoclonal antibody specific to the predominant TCR V-region product reverses paralytic disease²⁴. In addition, it is possible to prevent EAE by immunization with inactive encephalitogenic T cells²⁸ or with synthetic peptides from the CDR2 and CDR3 TCR V regions that are rearranged in the encephalitogenic clones^{29,30}. Elucidation of TCR expression in the brain may help in the design of similar treatments in MS patients. □

Received 8 December 1989; accepted 20 March 1990.

- McFarlin, D. E. & McFarland, H. F. *New Engl. J. Med.* **307**, 1183-1251 (1982).
- Raine, C. S. & Traugott, U. in *Immunoregulatory Processes in Experimental Allergic Encephalomyelitis and Multiple Sclerosis* (eds Alvord, E. A. & Kies, M. W.) 151-212 (Elsevier, New York, 1984).
- Prineas, J. W. & Wright, R. G. *Lab. Invest.* **38**, 409-421 (1978).
- Prineas, J. W. *Handbook of Clinical Neurology*, Vol. 3, 213-257 (Elsevier, New York, 1985).
- Woodroof, M. N. *et al. J. Neurol. Sci.* **74**, 135-152 (1986).
- Hafner, D. A. & Weiner, H. L. *Ann. Neurol.* **22**, 89-93 (1987).
- Hafner, D. A. & Weiner, H. L. *Immunol. Rev.* **100**, 307-332 (1987).
- Hoffman, F. M., Hinton, D. R., Johnson, K. & Merrill, J. E. *J. exp. Med.* **170**, 607-612 (1989).
- Hoffman, F. M. *et al. J. Immun.* **136**, 3239-3245 (1986).
- Traugott, U., Reinherz, E. L. & Raine, C. S. *J. Neuroimmun.* **4**, 201-221 (1983).

- Sobel, R. A. *et al. J. Immun.* **140**, 2210-2214 (1988).
- Hafner, D. A. *et al. J. exp. Med.* **167**, 1313-1322 (1988).
- Beall, S. S. *et al. J. Neuroimmun.* **21**, 59-66 (1989).
- Seboun, E. *et al. Cell* **57**, 1095-1100 (1989).
- Oksenberg, J. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 988-992 (1989).
- Saiki, R. K. *et al. Science* **230**, 1350-1354 (1985).
- Mullis, K. B. & Faloona, F. *Meth. Enzym.* **155**, 335-350 (1987).
- Sim, G. K. *et al. Nature* **312**, 771-775 (1984).
- Winoto, A. & Baltimore, D. *EMBO J.* **8**, 729-733 (1989).
- Loh, E. Y., Elliot, J. F., Cwiria, S., Lainer, L. L. & Davis, M. M. *Science* **243**, 217-220 (1989).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Yoshikai, Y., Kimura, N., Toyonaga, B. & Mak, T. W. *J. exp. Med.* **164**, 90-103 (1986).
- Klein, M. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 6884-6888 (1987).
- Acha-Orbea, H. *et al. Cell* **54**, 263-273 (1988).
- Urban, J. L. *et al. Cell* **54**, 577-582 (1988).
- Sakai, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 8608-8612 (1988).
- Heber Katz, E. & Acha-Orbea, H. *Immunol. Today* **10**, 164-169 (1989).
- Ben-Nun, A., Wekerle, H. & Cohen, I. R. *Nature* **292**, 60-61 (1981).
- Vandenbark, A. A., Hashim, G. & Offner, H. *Nature* **341**, 541-544 (1989).
- Howell, M. D. *et al. Science* **246**, 668-670 (1989).

ACKNOWLEDGEMENTS. This work was supported by grants from the NIH, the National Health and Medical Research Council (NH-MRC) of Australia, and the National Multiple Sclerosis Societies of the USA and Australia. J.R.O. is a fellow of the National Multiple Sclerosis Society (USA). R.B. is a fellow of the Alberta Heritage Foundation. C.C.A. is a Principal Research Fellow of the NH-MRC of Australia.

Nitric oxide as an inhibitory non-adrenergic non-cholinergic neurotransmitter

H. Buft, G. E. Boeckxstaens, P. A. Pelckmans, F. H. Jordaens, Y. M. Van Maercke & A. G. Herman

Divisions of Pharmacology and Gastroenterology, Faculty of Medicine, University of Antwerp (UIA), B-2610 Wilrijk, Belgium

INHIBITORY non-adrenergic non-cholinergic (NANC) nerves are thought to be important in the autonomic innervation of the gastrointestinal tract and other organ systems. The nature of their neurotransmitter is still debated. Speculation that nitric oxide (NO), formed from L-arginine in neuronal structures¹ and other cells², could act as a neurotransmitter, is not yet supported by demonstration of its release upon nerve stimulation. Using a superfusion bioassay, we report the release of a vasorelaxant factor upon stimulation of the NANC nerves³ in the canine ileocolonic junction. Several pieces of evidence, including the selectivity of the bioassay tissues, chemical instability, inactivation by superoxide anion and haemoglobin, inhibition by N^G-nitro-L-arginine (L-NNA)⁴ and potentiation by L-arginine all indicated that NO accounted for the biological activity of this transferable NANC factor.

The canine ileocolonic junction (ICJ) was isolated, superfused with Krebs-Ringer solution and the effluent superfused de-endothelialized rings of rabbit aorta⁵, arranged either in a cascade^{6,7} (Fig. 1) or in parallel. Upon electrical stimulation (ES; 16 Hz, 2 ms) or infusion of the nicotinic receptor agonist 1,1-dimethyl-4-phenylpiperazinium (DMPP), ICJ tissue released a vasodilator activity, causing 22.4 $\pm$ 3.2% (n = 12) and 26.3 $\pm$ 3.9% (n = 9) relaxation of the top tissue respectively. The release of the factor was frequency-dependent and its activity declined (39 $\pm$ 8%, n = 8, ES; 41 $\pm$ 15%, n = 4, DMPP) during passage down the cascade. As atropine and guanethidine were present and as the detector tissues failed to relax to acetylcholine (ACh) and noradrenaline (Figs 1 and 2), the transferable factor is indeed non-adrenergic and non-cholinergic. Furthermore, blockade of nerve conductance with tetrodotoxin abolished the release of the vasodepressant factor to both stimuli (Fig. 2), indicating that neuronal structures were activated. Non-selective stimulation of other cell populations² or microorganisms therefore seems unlikely.

The instability of the vasodilator factor was compatible with NO^{7,8}, as confirmed by the injection of authentic NO (Fig. 1). Moreover, haemoglobin (Hb), known to trap NO avidly⁷⁻⁹, eliminated the biological activity, whereas dilatation, due to NO

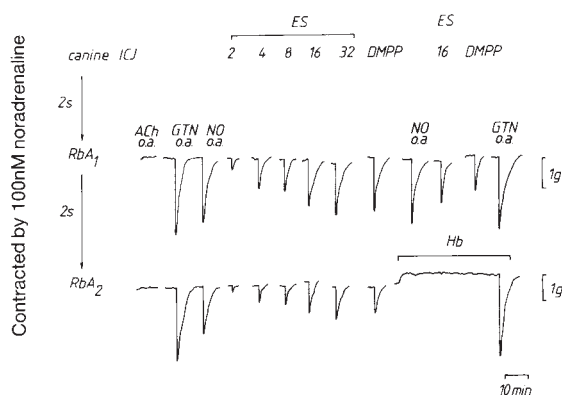


FIG. 1 The ICJ releases an unstable vasodilator in response to ES (2–32 Hz) or infusion of DMPP (30 μ M, 60 s, $n=4$). Biological activity of transferable factor and NO (0.05 nmol, injected over the aortae, o.a.) decreased to the same extent during the delay of 2 s between upper and lower aortic ring. Hb (100 nM) abolished both biological activities instantaneously, without affecting GTN (0.1 nmol, o.a.) responses. After removal of mucosa and submucosa, the ICJ tissue was placed in a superfusion chamber with platinum side walls, connected to a Grass stimulator (square pulses, 2 ms duration, 25 mA, 8 V for 20 s). The luminal side was superfused with Krebs-Ringer buffer (3 ml min⁻¹)⁵ containing SOD (20 units ml⁻¹), L-arginine (50 μ M) and guanethidine (3 μ M). The effluent superfused a cascade of two de-endothelialized rabbit aortic rings (RbA) contracted submaximally by 100 nM noradrenaline⁵. ACh (3 nmol, o.a.) injection confirmed the absence of endothelium. Thereafter atropine (300 nM) was introduced into the perfusate and the sensitivity of the bioassay was standardized by injection of GTN. Hb and NO solutions were prepared as described⁹.

formed from nitroglycerin (GTN) within the assay tissues, still occurred. Further evidence for the release of NO is provided by the findings that L-arginine, the precursor^{1,2} of NO, and superoxide dismutase (SOD), which opposes its inactivation by superoxide anion^{6,10}, enhanced the vasodepressant activity released from the ICJ (Fig. 2). The relaxations of the top aortic ring were augmented from $8.2 \pm 4.2\%$ to $24.2 \pm 4.4\%$ ($n=5$) by

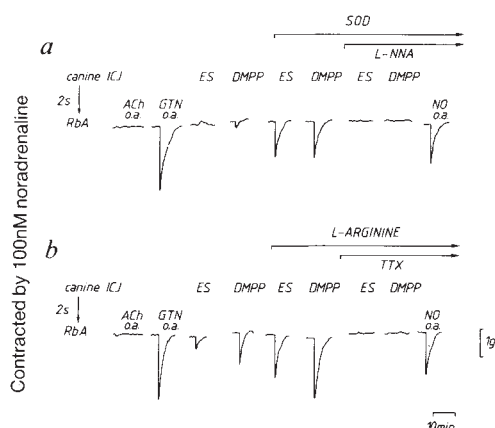


FIG. 2 Evidence for release of NO and its neuronal origin. SOD (20 U ml⁻¹, a) and L-arginine (50 μ M, b) enhance, whereas L-NNA (100 μ M, a) and tetrodotoxin (TTX, 1 μ M, b) abolish the release of the vasorelaxant factor in response to ES (16 Hz, 2 ms) and DMPP (30 μ M, 60 s). Relaxations induced by NO (0.05 nmol, o.a.) were not prevented. Effluent of two ICJ preparations (a and b) superfused separate aortic rings. L-Arginine was continuously present in (a), and SOD in (b). Contraction of assay tissues, confirmation of the absence of endothelium (ACh, 3 nmol, o.a.), atropine addition and calibration (GTN, 0.1 nmol, o.a.) were performed as described for Fig. 1.

SOD. Finally, L-NNA, an inhibitor of NO biosynthesis⁴ inhibited the release of the factor (Fig. 2).

Our data clearly demonstrate the release of NO during nerve stimulation. A role for NO in neurotransmission had already been postulated, and neuronal structures have the capacity to generate NO^{1,11}. Moreover, NO mimics NANC relaxations, blockers of NO biosynthesis inhibit these relaxations and L-arginine reverses this inhibition in the rat anococcygeus muscle^{12,13} and in the canine ICJ (G.E.B. *et al.*, submitted for publication). Our finding that NO is released upon nerve stimulation adds further evidence to propose NO or an NO-releasing molecule as inhibitory NANC neurotransmitter and extends the presence of the L-arginine: NO pathway² to gastrointestinal tissue. □

Received 3 January; accepted 13 March 1990.

- Knowles, R. G., Palacios, M., Palmer, R. M. J. & Moncada, S. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5159–5162 (1989).
- Moncada, S., Palmer, R. M. J. & Higgs, E. A. *Biochem. Pharmac.* **38**, 1709–1715 (1989).
- Pelckmans, P. A., Boeckxstaens, G. E., Van Maercke, Y. M., Herman, A. G. & Verbeuren, T. J. *Eur. J. Pharmac.* **170**, 235–242 (1989).
- Mülsch, A. & Busse, R. *Naunyn-Schmiedeberg's Archs. Pharmac.* **341**, 143–147 (1990).
- Bult, H., Fret, H. R. L., Van den Bossche, R. M. & Herman, A. G. *Br. J. Pharmac.* **95**, 1308–1314 (1988).
- Gryglewski, R. J., Palmer, R. M. J. & Moncada, S. *Nature* **320**, 454–456 (1986).
- Palmer, R. M. J., Ferrige, A. G. & Moncada, S. *Nature* **327**, 524–526 (1987).
- Ignarro, L. J., Buga, G. M., Wood, K. S., Byrns, R. E. & Chaudhuri, G. *Proc. natn. Acad. Sci. U.S.A.* **84**, 9265–9269 (1987).
- Kelm, M. *et al. Biochem. biophys. Res. Commun.* **154**, 236–244 (1988).
- Rubanyi, G. M. & Vanhoutte, P. M. *Am. J. Physiol.* **250**, H822–H827 (1986).
- Garthwaite, J., Charles, S. L. & Chess-Williams, R. *Nature* **336**, 385–388 (1988).
- Gillespie, J. S., Liu, X. & Martin, W. *Br. J. Pharmac.* **98**, 1080–1082 (1989).
- Ramagopal, M. V. & Leighton, H. J. *Eur. J. Pharmac.* **174**, 297–299 (1989).

Proton inhibition of *N*-methyl-D-aspartate receptors in cerebellar neurons

Stephen F. Traynelis & Stuart G. Cull-Candy

Department of Pharmacology, University College London, Gower Street, London, WC1E 6BT, UK.

MAMMALIAN neurons contain at least three types of excitatory amino-acid receptors, selectively activated by *N*-methyl-D-aspartate (NMDA) or aspartate, (S)- α -amino-3-hydroxy-5-methyl-4-isoxazole propionate ((S)-AMPA) and kainate¹. An important aspect of NMDA receptors is their regulation by a variety of factors such as glycine², Mg²⁺ (refs. 3, 4) and Zn²⁺ (refs 5, 6) that are present *in vivo*. We show here that NMDA receptor responses are selectively inhibited by protons, with a 50% inhibitory concentration (IC₅₀) that is close to physiological pH, implying that NMDA receptors are not fully active under normal conditions. (S)-AMPA and kainate responses remain unchanged at similar pH levels. Proton inhibition is voltage-insensitive and does not result either from fast channel block, a change in channel conductance, or an increase in the 50% excitatory concentration (EC₅₀) of aspartate/NMDA or glycine. Instead, protons seem to decrease markedly the opening frequency of 30–50 pS NMDA channels, and reduce the relative proportion of longer bursts. This feature of NMDA receptors could be relevant to neurotoxic activation of NMDA receptors during ischaemia⁷, as well as to seizure generation, as extracellular proton changes occur during both of these pathological situations^{8,9}. Furthermore, these results may have implications for normal NMDA receptor function as transient changes in extracellular protons occur during synaptic transmission^{8,10}.

Whole-cell currents produced by bath application of aspartate or NMDA (Fig. 1a) were reversibly decreased by $67.3 \pm 2.0\%$ ($\pm$ s.e.m.; $n=56$) when bath pH was reduced from 7.6 to 6.8; fast-membrane currents produced by brief ionophoretic pulses

of aspartate (Fig. 2a) were similarly decreased by $69.3 \pm 5.3\%$ ($n = 23$ cells). Alternatively, increasing the pH from 7.6 to 8.4 potentiated NMDA receptor currents by $55.3 \pm 6.8\%$ ($n = 39$). The calculated IC_{50} for the full inhibition curve (Fig. 1b) was pH 7.3, close to physiological pH and also similar to the pK_a values for free cysteine or histidine, suggesting that ionization of these amino-acid residues within the receptor might influence its function. In contrast, kainate and AMPA receptor-mediated currents were insensitive to protons over this pH range, but were inhibited by higher concentrations of H^+ , apparently in a non-competitive way ($n = 17$ –32 cells). We estimate that the IC_{50} for (S)-AMPA and kainate is pH 6.3 and 5.7, respectively (Fig. 1b). Several reports indicate that H^+ interacts with glutamate receptors^{14–17} (L. Vyklicky, V. Vlachova, and J. Krusek, personal communication).

There are many possible sites on the NMDA receptor at which protons might act. These include the binding sites for the agonists NMDA or aspartate, the strychnine-insensitive glycine-binding site², the Mg^{2+} -binding site^{3,4} and other sites within the channel. We have considered each of these possibilities in turn. Normalized whole-cell dose-response curves constructed at pH 6.8 or 7.6 for aspartate in the presence of 2–3 μM glycine, and also for glycine in the presence of 50 μM aspartate, are shown in Fig. 1c. There are no detectable differences in the EC_{50} calculated from the normalized curves for aspartate at pH 6.8 and 7.6, suggesting that proton inhibition is non-competitive with respect

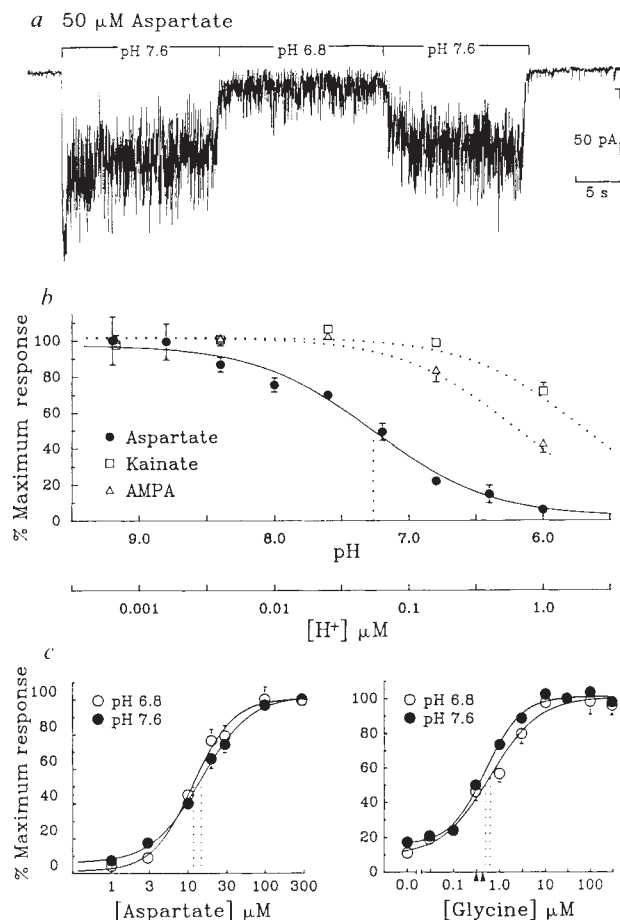
to the aspartate-binding site. Decreasing the pH from 7.6 to 6.8 increased the EC_{50} for glycine by 48.4%; however, the clear plateau of the dose-response curve in pH 6.8, and the inability of supplemental glycine (1 mM; $n = 6$ cells) to alter proton inhibition significantly, suggests that the effect of H^+ is non-competitive with respect to glycine. The suggestion¹⁶ that the increase in the estimated K_d for glycine at very low pH (6.3) contributes to proton inhibition through altered gating is consistent with the small shift we observed in the EC_{50} , but our data do not support competitive antagonism of glycine by H^+ over a more physiological pH range (6.8–7.6) in cerebellar granule cells.

Low concentrations of Mg^{2+} are known to block the NMDA receptor channel in a voltage-dependent manner^{3,4}. To test whether proton inhibition of NMDA receptor function is also voltage-sensitive, whole-cell currents were recorded in response to aspartate applied ionophoretically over a range of holding potentials at pH 6.8–8.4 ($n = 5$ –6 cells). As illustrated in Fig. 2b, aspartate currents reversed close to 0 mV and were linearly related to voltage between -80 to $+50$ mV in this pH range. The absence of additional rectification in the current-voltage plots at low pH suggests that proton inhibition does not involve a voltage-sensitive mechanism.

Because protons can influence cellular functions in many ways, we examined single-channel currents in various patch configurations¹⁸ to determine whether protons act directly on

FIG 1 a, The influence of protons on the whole-cell response to 50 μM aspartate plus 3 μM glycine applied by perfusion. During application, pH was stepped from 7.6 to 6.8, and back to 7.6 with no change in aspartate or glycine concentrations. b, Composite H^+ inhibition curve for NMDA receptor-mediated currents in response to 20–50 μM aspartate plus 3 μM glycine (filled circles). Values were calculated as a percentage of current at pH 7.6, and are shown normalized to 9.2 ($\pm$ s.e.m. when larger than symbol); measurements were taken from 3–33 cells at each pH. The data were fitted to the equation: $[(\text{maximum} - \text{minimum}) / (1 + ([H^+] / IC_{50}))] + \text{minimum}$. The calculated IC_{50} (dotted line) for aspartate was pH 7.3 (50 nM H^+). Whole-cell currents to 30–50 μM kainate ($n = 4$ –25 cells; open squares) and 10–15 μM (S)-AMPA ($n = 4$ –11 cells; open triangles) at pH 9.2, 8.4, 7.6, 6.8, and 6.0 are included. Aspartate ($pK_{a-NH_2} = 9.7$) concentrations were adjusted appropriately at pH 8.8 and 9.2, assuming that only the fully ionized species is active. c, Dose-response curves for aspartate and glycine at pH 6.8 and 7.6 were normalized to the maximum response, and fitted to the equation $[(\text{maximum} - \text{minimum}) / (1 + (EC_{50} / [\text{aspartate}])^n)] + \text{minimum}$, where n is the Hill coefficient; on average, it was 1.52 for aspartate and 0.70 for glycine. The calculated EC_{50} values to varied aspartate concentration plus 3 μM glycine at pH 6.8 ($n = 3$ –10 cells) and 7.6 ($n = 3$ –12 cells) were 12 and 14 μM , respectively (dotted lines); responses to variable glycine plus 50 μM aspartate had EC_{50} values (arrows) of 0.46 μM in pH 6.8 ($n = 4$ –11 cells) and 0.31 μM in pH 7.6 ($n = 4$ –11) when the minimum was constrained to 0%²⁶; dotted lines indicate EC_{50} values calculated with a variable minimum. Values were expressed as a percentage of the mean of control and recovery, and averaged throughout the experiment to compensate for receptor rundown²⁷.

METHODS. Cerebellar granule cell explants were prepared as previously described²⁰. Recordings were made at 22–25 °C from identified granule cells²⁰. The bathing solution contained (mM): 150 NaCl; 2.8 KCl; 1.0 $CaCl_2$; 10.0 sodium HEPES. If drugs were applied at different pH, the cell was first washed at the new pH to inactivate the transient proton-activated Na^+ current²⁸. Our reductions in pH do not significantly alter the ionization of free aspartate or glycine. For whole-cell and outside-out experiments the pipette solution (intracellular) contained (mM): 140 CsCl (or 110 CsF, 30 CsCl); 4 NaCl; 5 potassium EGTA; 5 HEPES; 0.5 $CaCl_2$ pH 7.2; Pipette solutions for cell-attached and inside-out patch experiments consisted of (mM): 150 NaCl; 2.8 KCl; 1.0 $CaCl_2$; 10 HEPES plus 30–50 μM aspartate, 3–5 μM glycine at pH 7.2–7.6. Open and shut durations were measured by the time-course fitting method²⁹. Bursts were distinguished using a t_{crit} (mean 1.2 ms) calculated between the two brief ($\tau_1 = 67$, $\tau_2 = 537 \mu s$) and two longer ($\tau_3 = 23.4$, $\tau_4 = 524$ ms; $n = 10$ patches) exponential components describing shut time distributions at pH 7.6. Open probability was calculated from patches subjected to both pH 6.8 and 7.6 using fitted single openings when double openings were $< 0.5\%$ of total, otherwise determined from the



integrated records. The open probability in patches used for burst analysis was, on average, 0.0126 and 0.0039 at pH 7.6 and 6.8, respectively ($n = 8$); fitted parameters were similar in patches with a high $P(\text{open})$ compared with low $P(\text{open})$. In cell-attached patches, the major aspartate-activated conductance levels were determined to be 40–50 pS; frequent transitions were observed between the 40 and 50 pS levels.

the NMDA receptor-channel complex or, alternatively, through an indirect mechanism. In outside-out patches, where most intracellular constituents are assumed to have been removed, the probability of an aspartate-activated channel (30, 40, 50 pS; glycine present) being open was decreased by $63.6 \pm 6.2\%$ when the pH was reduced from 7.6 to 6.8 ($n=9$ patches; Fig. 2d). Furthermore, when cells were subjected to a reduced pH that would block the whole-cell response to aspartate (pH 6.8), single-channel openings produced by aspartate in the pipette could still be recorded in cell-attached patches, and the probability that they were open was not significantly different from that in pH 7.6 ($n=7$ patches; paired *t*-test; power to detect a 60% change was 0.75). Similarly, the probability of aspartate-activated channels being open was not significantly different when the intracellular surface of inside-out patches was exposed to pH 6.8 or 7.6 ($n=5$ patches; power to detect a 40% change was 0.75). These experiments provide strong support for the idea that H^+ acts at a specific site localized on or near the extracellular face of the NMDA receptor-channel complex.

Although the voltage insensitivity of H^+ inhibition meant that a channel-blocking mechanism was unlikely, it was possible that H^+ produced an incomplete blockage of channel currents through actions at a site outside the membrane electric field, as reported for proton block of Ca^{2+} channels¹⁹. This could give rise to discrete lower conductance levels at reduced as well as normal pH values. We therefore examined the effect of H^+ on both aspartate-activated single-channel currents in outside-out

patches and aspartate or NMDA-induced whole-cell noise. The single-channel conductance estimated from aspartate and NMDA noise spectra fitted with two Lorentzian components²⁰ was unchanged at both inhibiting and potentiating pH (53 pS at pH 7.6; $n=33$ spectra in 11 cells; single-factor analysis of variance). Consistent with this result, single-channel conductance levels of approximately 11, 19, 28, 43 or 53 pS (refs 21, 22) were all present in outside-out patches exposed to aspartate at pH 6.8 (Fig. 2c), and the relative frequencies of single-channel openings to the 19, 28, 43, 53 pS (refs 20, 23) conductance levels were similar to their counterparts determined at pH 7.6 ($n=10$ patches).

Analysis of open times (50–85 μ s resolution) revealed a decrease of $54.8 \pm 10.9\%$ ($n=10$ patches) for the burst frequency (corrected for missed events). The decreased burst frequency did not appear to reflect enhanced desensitization, as the ratio of peak to steady-state somatic currents was insensitive to the pH (glycine present; $n=10$ cells). The mean duration of all bursts was also decreased by $38.2 \pm 7.7\%$ ($n=10$). Taken together, these measured decreases in burst frequency and mean burst length at pH 6.8 would be expected to cause a 66.1% reduction in whole-cell current, and therefore would fully account for our whole-cell measurements. To examine the effect of H^+ on burst characteristics, we compared the three fitted components²⁰ that described distributions both of burst lengths and of individual open times at pH 7.6 and 6.8. Protons did not significantly change the fitted time constants describing either

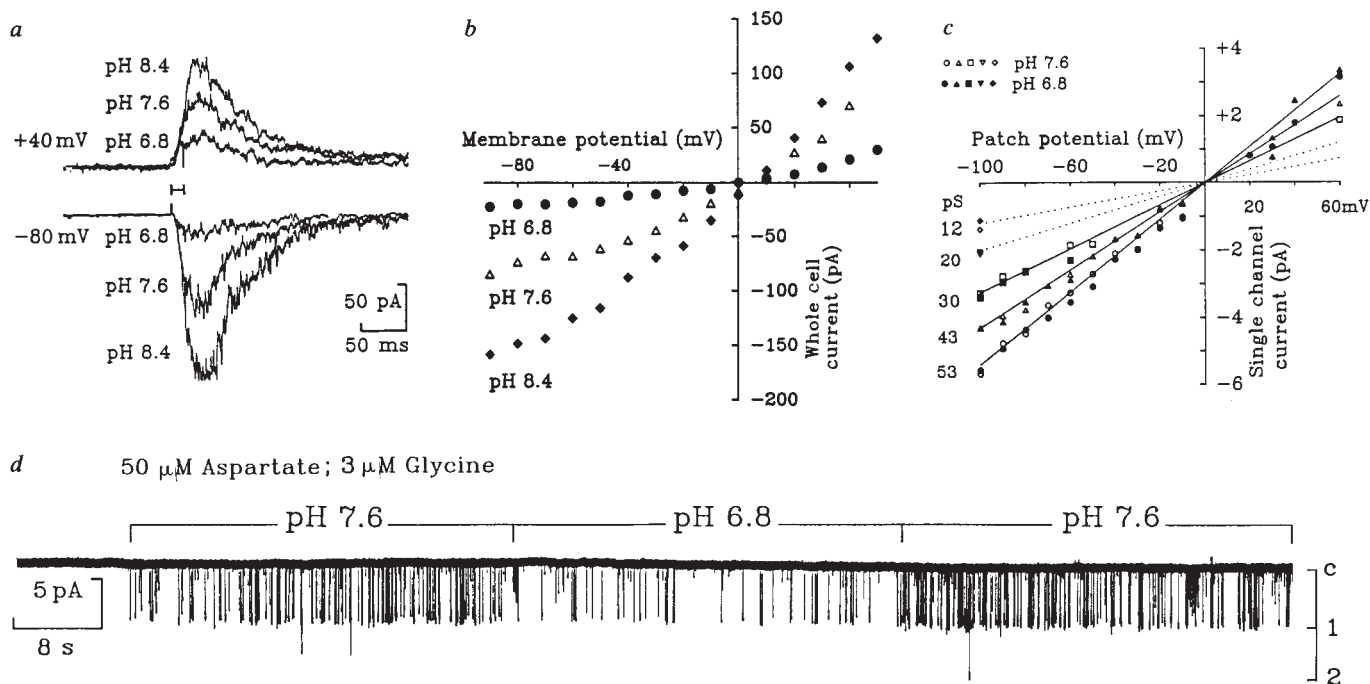


FIG. 2 *a*, Whole-cell responses to brief (10–30 ms) ionophoretically applied aspartate (indicated by bar; 100–250 nA; pipette resistance 90–130 M Ω) in a cell at +40 and –80 mV bathed in pH 6.8, 7.6, and 8.4 (3 μ M glycine in the bath). Whole-cell currents were digitized at 1 kHz (filtered 0.5 kHz, –3 dB; calibration 50 pA, 50 ms), averaged (3–5), and 100 ms of the peak current measured. *b*, Whole-cell current-voltage relationship in response to ionophoretically applied aspartate in nominally Mg^{2+} -free solution (plus 3 μ M glycine) at pH 6.8, 7.6, 8.4. *c*, Single-channel current-voltage relationship for multiple conductance levels produced by aspartate at pH 6.8 ($n=4$) and 7.6 ($n=5$) over the range –100 to +80 mV. Current amplitudes were extracted from amplitude histograms constructed from the digitized data selected for channel open times; histograms were fitted to the sum of 1–5 Gaussian densities²¹; 50–1000 openings were analysed at each potential. Low-variance mean amplitude histograms³⁰ were constructed to aid in sublevel detection. Mean current amplitudes for the ~12, 20 pS conductance

levels are only shown at –100 mV for clarity. Slope conductances calculated (linear regression) from the averaged currents at pH 7.6 were 53.0, 42.5, 31.6, 21.1, 12.2 pS; and at pH 6.8 were 53.4, 42.6, 27.6, 19.2, 11.4 pS. Lines are for averaged slopes at pH 7.6 and 6.8, constrained through the origin. The average relative opening frequencies ($n=10$ patches) for the multiple conductance levels at –80 mV in pH 6.8, 7.6 were 0.65, 0.76 (53 pS); 0.19, 0.15 (43 pS); 0.03, 0.03 (30 pS); 0.04, 0.03 (20 pS); 0.09, 0.03 (12 pS). *d*, Aspartate-induced single-channel currents in an outside-out patch at –80 mV, recorded on slow time-base; calibration 5 pA, 8 s. 50 μ M aspartate plus 3 μ M glycine was applied externally at pH 7.6, and the pH stepped to 6.8 without a change in agonist concentration, and back to 7.6 (as indicated). The probability of being open for this patch was 0.0121 in pH 7.6, and 0.0037 in pH 6.8. The closed level (c) and amplitude of 1, 2 openings are indicated. These data were filtered at 2.5 kHz (–3 dB).

of these distributions, for which the burst lengths at pH 7.6 were 0.05 ± 0.01 , 1.21 ± 0.33 , and 6.13 ± 0.51 ms ($n=8$ patches). Instead, the decrease in overall mean burst length in patches at pH 7.6 and subsequently at pH 6.8 reflected a decrease of $42.9 \pm 9.9\%$ ($n=7$ patches; $P<0.05$; paired t -test) in the relative fitted area of the longest component. Thus, although the time constants describing burst distributions did not change, there seemed to be a smaller proportion of long bursts. Analogous results were found in individual open time histograms (data not shown).

Substantial decreases in extracellular pH (~ 0.6 units^{10,24}) are known to accompany increased extracellular glutamate levels⁷ during experimental ischaemia, and they might attenuate neurotoxic activation of NMDA receptors⁷. This possibility is supported by findings that reduced pH protects cultured neurons from glutamate-induced and hypoxia-induced toxicity²⁵. Furthermore, the activation of synaptic NMDA receptors contributes to seizure generation in most experimental situations⁷, and the biphasic changes in extracellular H^+ reported for stimulus-induced seizures *in vivo*⁹ could influence synaptic NMDA receptors, and therefore seizure time course. Indeed, this idea is consistent with recent findings that small changes in H^+ (± 0.1 pH unit) alter neuronal excitability and modulate seizure threshold⁹.

The proton sensitivity of NMDA receptors could also have important implications for normal neuronal function. Extracellular pH is known to change by up to 0.4 pH units during electrical stimulation of cerebellar or hippocampal slices^{8,11}, or of cerebellar parallel fibres *in vivo*¹⁰. Moreover, from our results, acidic changes of such magnitude should inhibit NMDA receptors by $45.9 \pm 9.2\%$ ($n=7$ cells). Of particular interest is the possibility that the fast (~ 10 ms) acid transient previously reported¹¹ might arise from acidification of the cleft by the contents of synaptic vesicles, as an H^+ -ATPase provides the electrochemical H^+ gradient (inside is acidic and positively charged) that drives vesicular glutamate uptake^{12,13}. This would inhibit NMDA receptors, but not kainate or AMPA receptors, and add another level of complexity to synaptic events. □

Migration of myoblasts across basal lamina during skeletal muscle development

Simon M. Hughes & Helen M. Blau

Department of Pharmacology, Stanford University School of Medicine, Stanford, California 94305-5332, USA

BASAL lamina is a sheet of extracellular matrix that separates cells into topologically distinct groups during morphogenesis and is thought to form a barrier to cell migration. We have examined whether, during normal muscle development, myoblasts—mononucleate muscle precursor cells¹—can cross the basal lamina that surrounds each multinucleate muscle fibre². We marked myoblasts *in vivo* by injecting replication-defective retroviral vectors encoding LacZ into muscle tissue and analysed the fate of their progeny by the expression of β -galactosidase^{3,4}. A dual labelling method with broad application to retroviral lineage-marking studies was developed to ensure that most clusters of labelled cells were clones derived from a single precursor cell. Most of the myoblasts that were infected at a late stage of rat hindlimb development, when each fibre with its satellite myoblasts is individually encased in a basal lamina sheath^{2,5,6}, gave rise to clones that contributed to several labelled fibres. Our results show that myoblasts from healthy fibres migrate across basal lamina during normal development and could contribute to the repair of fibres damaged by injury or disease.

To mark single myoblasts and their progeny, we injected rat skeletal muscles with the recombinant BAG retroviral vector³. This replication-defective retrovirus is a useful marker of cell lineage, as after infection it stably integrates into the host genome, expresses β -galactosidase (β -gal) and is inherited by all progeny cells at mitosis. A histochemical analysis of serial cryostat sections of hindlimbs 2 weeks after infection revealed muscles fibres labelled with β -gal distributed in regions

TABLE 1 β -gal-labelled muscle fibres were frequently in clusters of more than one fibre

Experiment	Age (days) at		Regions with	
	Injection	Analysis	One fibre	Multiple fibres
1	9	23	9	16
2	14	38	9	10
3	14	38	5	5
4	15	37	16	22
5	16	29	1	1
6	16	29	1	3
7	19	35	0	1
Total			41 (41%)	58 (59%)

Retroviral vector, prepared as described in the legend to Fig. 1 was injected into rat hindlimbs on or after P9, a developmental stage when the final adult number of muscle fibres has been formed and further growth of muscle is due to an increase in size of individual fibres, largely by myoblast fusion into pre-existing fibres. Three observations ensured that in our experiments new fibres were not being formed: (1) counts of muscle fibre number were similar at P9 and P30. For soleus, they were $2,547 \pm 233$ ($n=3$) and $2,357 \pm 213$ ($n=3$), respectively; (2) very few, if any, small-diameter nascent fibres were present after P9; (3) the greatest diameter of fibres within multifibre clones, single-fibre clones and unlabelled adjacent fibres did not differ significantly, and was $25 \pm 4 \mu\text{m}$ ($n=14$), $28 \pm 4 \mu\text{m}$ ($n=4$) and $24 \pm 6 \mu\text{m}$ ($n=94$), respectively (mean $\pm$ s.d.) at P23 (experiment 1). Serial transverse sections indicated that labelled myofibres do not bifurcate, in agreement with electron microscopic studies^{13,24}. Regions of labelled fibres contained one or more fibres separated from one another by less than three fibre diameters and from all other labelled fibres by at least ten fibre diameters. In experiments 2 and 3 (also shown in Table 2), a mixture of BAG and MMuLVSVnLacZ was used.

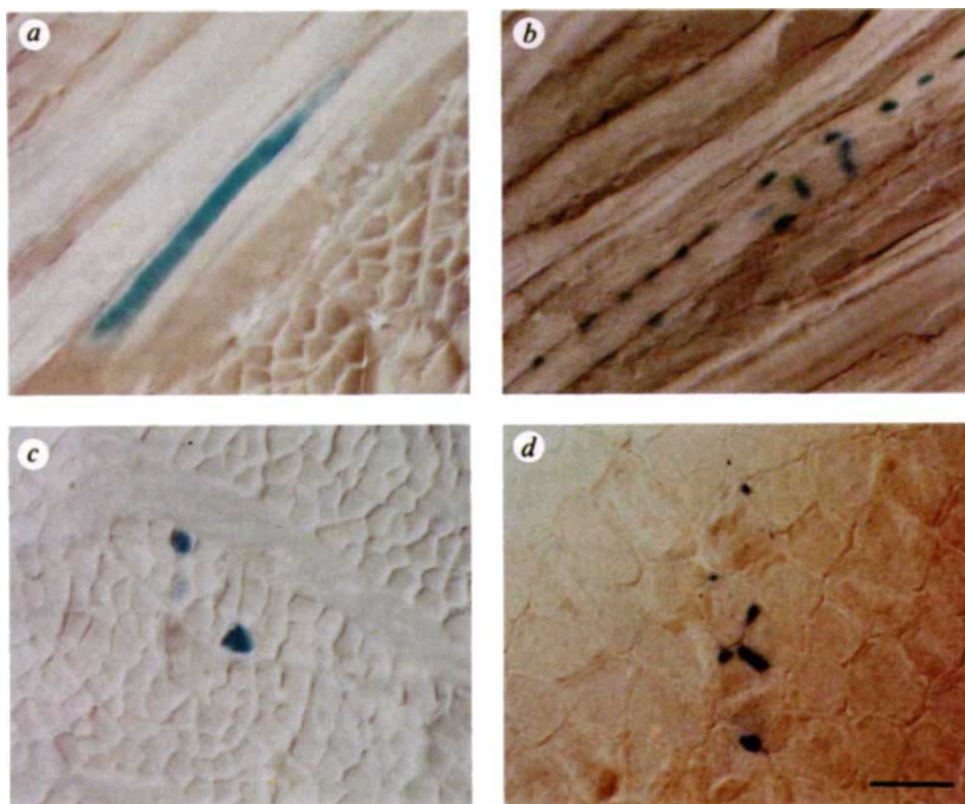
Received 8 December 1989; accepted 2 March 1990.

- Watkins, J. C. & Evans, R. H. A. *Rev. Pharmacol. Toxicol.* **21**, 165–204 (1981).
- Johnson, J. W. & Ascher, P. *Nature* **325**, 529–531 (1987).
- Mayer, M. L., Westbrook, G. L. & Guthrie, P. B. *Nature* **309**, 261–263 (1984).
- Nowak, L., Bregestovski, P., Ascher, P., Herbet, A. & Prochiantz, A. *Nature* **307**, 462–465 (1984).
- Peters, S., Koh, J. & Choi, D. W. *Science* **236**, 589–593 (1987).
- Mayer, M. L., Vyklicky, L. & Westbrook, G. L. *J. Physiol.* **415**, 329–350 (1989).
- Dingledine, R. et al. *CRC Critical Rev. Neurobiol.* **4**, 1–96 (1988).
- Chesler, M. & Chan, C. Y. *Neuroscience* **27**, 941–948 (1988).
- Balestrino, M. & Somjen, G. G. *J. Physiol. Lond.* **396**, 247–266 (1988).
- Kraig, R. P., Ferreira-Filho, C. R. & Nicholson, C. *J. Neurophysiol.* **49**, 831–850 (1983).
- Krishtal, O. A., Osipchuk, Y. V., Shelest, T. N. & Smirnov, S. V. *Brain Res.* **436**, 352–356 (1987).
- Maycox, P. R., Deckwerth, T., Hell, J. W. & Jahn, R. *J. Biol. Chem.* **263**, 15423–15428 (1988).
- Cidon, S. & Sihra, T. S. *J. Biol. Chem.* **264**, 8281–8288 (1989).
- Grantyn, R. & Lux, H. D. *Neurosci. Lett.* **89**, 198–203 (1989).
- Gruol, D. L., Barker, J. L., Huang, L.-Y. M., MacDonald, J. F. & Smith, T. G. *Brain Res.* **183**, 247–252 (1980).
- Tang, C. M., Dichter, M. & Morad, M. *Soc. Neurosci. Abstr.* **15**, 326 (1989).
- Morad, M., Dichter, M. & Tang, C. M. *Soc. Neurosci. Abstr.* **14**, 791 (1988).
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Archiv.* **391**, 85–100 (1981).
- Pietrobon, D., Prodhon, B. & Hess, P. *Nature* **333**, 373–376 (1988).
- Cull-Candy, S. G., Howe, J. R. & Ogden, D. C. *J. Physiol., Lond.* **400**, 189–222 (1988).
- Cull-Candy, S. G. & Usowicz, M. M. *Nature* **325**, 525–528 (1987).
- Jahr, C. E. & Stevens, C. F. *Nature* **325**, 522–525 (1987).
- Howe, J. R., Colquhoun, D. & Cull-Candy, S. G. *Proc. R. Soc. B233* 407–422 (1988).
- Mutch, W. A. C. & Hansen, A. J. *J. Cereb. Blood Flow Metab.* **4**, 17–27 (1984).
- Giffard, R. G., Monyer, H., Christine, C. W. & Choi, D. W. *Brain Res.* **506**, 339–342 (1990).
- Kieckner, N. W. & Dingledine, R. *Science* **241**, 835–837 (1988).
- MacDonald, J. F., Mody, I. & Salter, M. W. *J. Physiol.* **414**, 17–34 (1989).
- Krishtal, O. A. & Pidoplichko, V. I. *Neuroscience* **5**, 2325–2327 (1980).
- Colquhoun, D. & Sigworth, F. J. in *Single Channel Recording* (eds Sakmann, B. & Neher, E. 191–263 (New York, Plenum, 1983).
- Pattlak, J. B. *J. gen. Physiol.* **92**, 413–430 (1988).

ACKNOWLEDGEMENTS. This work was supported by the Wellcome Trust. We thank Corinne Symonds and Jackie Carr for assistance with tissue culture, and David Colquhoun, Alasdair Gibb, and Nancy Mulrine for helpful discussions and comments on the manuscript. S.F.T. is in receipt of an NIH Fellowship.

FIG. 1 Retroviral vectors target β -galactosidase to the cytoplasm or nucleus of muscle fibres *in vivo*. BAG, which expresses cytoplasmic β -galactosidase (a, c), or MMuLVSVnlsLacZ, which expresses nuclear β -galactosidase (b, d), was injected into rat hindlimb muscle at postnatal day 5 (P5) (a, c), P1 (b) or P15 (d) and analysed by serial section on P18 (a, longitudinal section; c, transverse section), P21 (b, longitudinal section) or P37 (d, transverse section), respectively. Between P21 and P37 muscle fibres increased threefold in diameter (compare c and d). Apparent nuclear diameter (b, d) varies with the plane of section. Scale bar, 50 μ m. The cluster in c was in the lateral gastrocnemius separated by 40 fibre diameters from its closest labelled neighbour (a) in the biceps femoris.

METHODS. Supernatant from confluent CREBAG α or Ψ -2-21-C₃ retroviral producer cells made by transfecting BAG into the Ψ _{CRE} or pMMuLVSVnlsLacZ into the Ψ ₂ ecotropic packaging cell lines respectively was collected, filtered, and concentrated by centrifugation. Wild-type helper virus was absent. Viral stock (50–200 μ l) with 10 μ g ml⁻¹ polybrene was mixed with charcoal particles to locate the site of injection and administered as a single injection from a 26-gauge needle into the latero-dorsal surface of the hindlimb of each anaesthetized Wistar rat. After 2–3 weeks, animals were analysed for the expression of β -gal. Muscle was fixed by cardiac perfusion with 4% paraformaldehyde, 0.5% glutaraldehyde, 100 mM PIPES buffer pH 7.4. After 30–60 min, lower hindlimbs were dissected free of skin and placed in the same fixative overnight at 4 °C, then in 30% sucrose PBS for 24 h at 4 °C, frozen in freezing isopentane and cut into serial 30 μ m sections on a Bright cryostat. In some experiments, sections were cut on unfixed tissue. All sections were postfixed in 4% paraformaldehyde in PBS washed and stained in 1 mg ml⁻¹ 5-bromo-4-chloro-3-indonyl-D-galactoside (X-gal), 5 or 35 mM potassium ferri- and ferrocyanide, 1 mM MgCl₂ in PBS overnight at 30 °C, mounted in glycerol-PBS (9:1) and examined under bright-field optics on a Zeiss Axiophot for the presence of the blue X-gal reaction product. Charcoal particles were



most frequently located between soleus and lateral gastrocnemius. The nuclear localization of nls- β -gal in tissue sections was confirmed by counter-staining with Hoechst 33258. Like cyt- β -gal, nls- β -gal diffuses within fibres and therefore labels nuclei in addition to those that encoded it¹⁰. Tandem arrays of 200 labelled nuclei were observed in single fibres *in vivo* (b). No detectable differences in frequency, morphology, size or distribution of clusters formed by BAG or MMuLVSVnlsLacZ were seen, consistent with the expectation that these vectors differ only in the distribution of β -gal to nuclear and cytoplasmic cell compartments. Clusters close to charcoal particles did not differ from those several millimetres away, suggesting that the injection did not perturb development of nearby tissue.

throughout the limb (Fig. 1), indicating that the retroviral vector could gain access to and infect muscle cells *in vivo*. Endogenous β -gal activity was negligible in muscle, as the muscles of uninjected or mock-injected animals had no labelled muscle fibres. Regions containing both labelled muscle and nonmuscle cells were not seen (data not shown), consistent with the distinct somitic origin of limb muscle in birds^{7,8}. As retroviruses only infect dividing cells⁹, the muscle fibres that expressed β -gal presumably resulted from infection of mitotic myoblasts and

the subsequent fusion of their clonal progeny into fibres.

Labelled muscle fibres occurred either alone or in small clusters. Figure 1a shows a typical longitudinal section of a single labelled muscle fibre that contained β -gal detectable over a length of about 100 nuclei, equivalent to 2 mm. This finding indicated that β -gal is free to diffuse along muscle fibres^{10,11}, in contrast to some endogenous muscle gene products¹² and, as shown in the cluster of three fibres in Fig. 1c, can be analysed in transverse sections. Elsewhere in this leg, 20 regions of one

TABLE 2 Cytoplasmic and nuclear β -gal-labelled fibres segregate into distinct clusters

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Total
Experiment 1																				
Nls- β -gal fibres (n)	0	0	0	0	0	0	0	0	1	0	0	2	0	0	0	0	4	0	0	7
Cyt- β -gal fibres (n)	1	1	1	1	1	1	1	1	0	2	2	0	3	3	3	4	0	5	6	36
Experiment 2																				
Nls- β -gal fibres (n)	0	0	0	0	0	0	0	2	0	0	0									2
Cyt- β -gal fibres (n)	1	1	1	1	1	1	2	0	3	3	14									26

Injection of a mixture of BAG and MMuLVSVnlsLacZ retroviral vectors into rat hindlimbs revealed that fibres within clusters are generally labelled by the same vector (χ^2 analysis: $P > 0.01$, although this value is limited by the small cluster sizes). Limbs were injected at P14 with a 6:1 mixture (BAG:MMuLVSVnlsLacZ) as described in Fig. 1 and analysed in serial sections ~3 weeks later (P38).

or more labelled fibres were distributed in five muscles and separated from their nearest neighbouring region of labelled fibres by 0.47 ± 0.25 mm. Therefore clusters of labelled fibres were generally well-separated and distributed among many different muscles of the limb.

To determine whether myoblasts migrate across basal lamina, retroviral vector was injected on or after postnatal day 9 (P9). At this stage, unlike the early postnatal period (P0–P5) when myoblasts encased in the basal lamina of one fibre give rise to new fibres within that lamina^{2,6}, each myoblast is encased in a basal lamina in close association with a single fibre, the number of fibres typical of adult muscle has been reached^{5,6}, and the growth of muscle tissue occurs by the fusion of myoblasts into pre-existing fibres. An analysis of serial sections of hindlimb muscles injected on or after P9 in seven independent experiments revealed that 59% of β -gal-positive regions contained more than one labelled fibre (Table 1). Because basal lamina separated each fibre and its myoblasts from its neighbours, clusters of labelled fibres were likely to be due to the migration of myoblasts from one fibre to another.

The interpretation that myoblasts migrate relies on several labelled fibres within a cluster arising from infection of a single myoblast. Alternatively, each labelled fibre within a cluster could have resulted from a separate infection event. To rule out the latter possibility we performed two types of experiments. First, the retroviral vector was titrated so that clusters were separated by greater distances. Multifibre clusters still occurred at a high frequency (Table 1, experiments 5–7). Second, we injected a mixture of two different retroviral vectors: BAG, which expresses β -gal in the cytoplasm (cyt- β -gal) and MMuLVSVnlsLacZ (ref. 4), which targets β -gal to the nucleus because of an added N-terminal SV40 T-antigen nuclear localization sequence (nls- β -gal). These two vectors possess the same *env*-encoded protein and therefore have similar infectivity⁴, yet are readily distinguishable *in vivo* (Fig. 1, compare *a, c* with *b, d*). Table 2 shows that fibres within a cluster contained either nls- β -gal or cyt- β -gal but not both. Therefore similarly labelled fibres were generally associated with one another in clusters. Given the observed cluster sizes, the probability of this association occurring if each labelled fibre arose from a separate infection is $<1/10^7$. We conclude that a cluster of similarly labelled fibres arose by infection of a single myoblast, migration of its progeny across basal lamina, and fusion into adjacent fibres.

Although the data in Table 2 strongly indicate that each cluster represents a clone, the number of clusters that could be analysed was relatively low, as most muscle growth had occurred before infection. To determine the proportion of clusters that represent single myogenic clones, we performed the dual vector experiment in rats injected immediately after birth (P0) (Table 3). At this early stage of development the rate of myoblast division is high and new fibres are still being formed, leading to a relatively high frequency of infection and formation of many large multifibre clones. Therefore, the chance of obtaining overlapping clones is greater at this stage, providing a rigorous test of whether clusters are clonal. Serial sections of two hindlimbs revealed that of 87 multifibre clusters analysed, only two clusters ($<3\%$) contained adjacent fibres labelled with nls- β -gal and cyt- β -gal (Table 3). After correction for the probability of a double infection by the same type of vector, these data show that $91 \pm 8\%$ ($P > 0.95$) of clusters were clonal in animals infected at a titre that resulted in ~ 40 clusters of labelled fibres per limb. This dual retroviral vector technique should prove generally useful in lineage studies for determining the frequency with which clusters of labelled cells represent clones.

Myoblast migration across basal lamina is not a rare event. The chance that clusters were clones in animals at late stages of development when fibre formation was complete (Tables 1 and 2) is likely to be even greater than 91%, because the frequency of clusters per limb was lower (<20 per limb). As 59% of labelled regions contained multifibre clusters at this

TABLE 3 Most clusters of β -gal-labelled fibres are clones

Experiment	Number of multifibre clusters with		
	Nls- β -gal	Cyt- β -gal	Cyt- β -gal and nls- β -gal
1	34	11	2
2	30	10	0
Total	64 (74%)	21 (24%)	2 (2%)

Injection of a mixture of retroviral vectors at P0 and analysis at P23, when clusters of large size are obtained with high frequency, revealed that most clusters of β -gal-labelled fibres are clones. The proportion of clusters that contained both nls- β -gal (MMuLVSVnlsLacZ) and cyt- β -gal (BAG) was approximately 2%. On the basis of this value, the predicted frequencies of two infections by the same type of vector were calculated as binomial estimates and were 6% for nls- β -gal and 1% for cyt- β -gal. By adding these estimates to the observed cyt- β -gal and nls- β -gal double infection value (2%) the frequency with which regions of labelled fibres derive from single infection events and are clonal approximates $91 \pm 8\%$ ($P > 0.05$). This value corrects for the ratio of the two vectors. The error was calculated by combining the binomial variance of each estimate.

stage, we conclude that at least 50% of infected myoblasts gave rise to progeny that migrated across basal lamina and fused into neighbouring myofibres. Moreover, myoblast migration could encompass at least five separate fibres (Table 2). Whether myoblasts can migrate further remains to be determined.

Our studies indicate that, during normal development, myoblasts frequently migrate from one muscle fibre to another at a time when the tissue is divided into a series of compartments by basal lamina. This finding is consistent with electron micrographs of undamaged muscle showing cells extending through basal lamina^{13,14} and suggests that myogenic cells are present in the muscle interstitium. Previous studies suggested that myoblasts migrate^{15–18}, but it was unclear whether the myoblasts could traverse intact basal lamina, because in each study the tissue was traumatized. This could have led to invasion of leukocytes, which in addition to tumour cells are one of the few cell types known to cross basal lamina^{19–21}, that provided a migration route for myoblasts. It remains to be determined whether diffusible cues²² or the composition of the basal lamina²³ regulate the migration of myoblasts *in vivo*, perhaps recruiting them to sites of injury.

That myoblasts can migrate from one healthy fibre to another lends promise to proposed therapies for Duchenne muscular dystrophy (DMD) based on myoblast implantation¹⁸. In degenerating muscle, it is possible that myoblasts migrate greater distances than those observed here. If human myoblasts, like rat myoblasts, can migrate across basal lamina, then injection of myoblasts into a few sites could constitute an efficient therapy for large regions of diseased muscle. □

Received 24 January 1990; accepted 5 March 1990.

1. Mauro, A. *J. biophys. biochem. Cytol.* **9**, 493–494 (1961).
2. Kelly, A. M. & Zacks, S. I. *J. Cell Biol.* **42**, 135–153 (1969).
3. Price, J., Turner, D. & Cepko, C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 156–160 (1987).
4. Bonnerot, C., Rocancourt, D., Briand, P., Grimmer, G. & Nicolas, J. F. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6795–6799 (1987).
5. Enesco, M. & Puddy, D. *Am. J. Anat.* **114**, 235–244 (1964).
6. Ontell, M. & Kozeka, K. *Am. J. Anat.* **171**, 133–148 (1984).
7. Chevallier, A., Kieny, M. & Mauger, A. *J. Embryol. exp. Morphol.* **41**, 245–258 (1977).
8. Christ, B., Jacob, H. J. & Jacob, M. *Anat. Embryol. Berl.* **150**, 171–186 (1977).
9. Varmus, H. E., Padgett, T., Heasley, S., Simon, G. & Bishop, J. M. *Cell* **11**, 307–319 (1977).
10. Ralston, E. & Hall, Z. W. *Science* **244**, 1066–1069 (1989).
11. Hughes, S. M. & Blau, H. M. in *The Dynamic State of Muscle Fibres*, (ed. Pette, D.) (Walter de Gruyter & Co., Berlin–New York, in the press).
12. Pavlath, G. K., Rich, K., Webster, S. G. & Blau, H. M. *Nature* **337**, 570–573 (1989).
13. Ontell, M. *Anat. Rec.* **189**, 669–690 (1977).
14. Kelly, A. M. *Anat. Rec.* **190**, 891–903 (1978).
15. Schultz, E., Jaryszak, D. L., Gibson, M. C. & Albright, D. J. *J. Muscle Res. Cell Motil.* **7**, 361–367 (1986).
16. Watt, D. J., Morgan, J. E., Clifford, M. A. & Partridge, L. A. *Anat. Embryol. Berl.* **175**, 527–536 (1987).
17. Lipton, B. H. & Schultz, E. *Science* **205**, 1292–1294 (1979).
18. Partridge, T. A., Morgan, J. E., Coulton, G. R., Hoffman, E. P. & Kunkel, L. M. *Nature* **337**, 176–179 (1989).
19. Anderson, A. O. & Anderson, N. D. *Immunology* **31**, 731–748 (1976).

20. Warfel, K. A. & Hull, M. T. *Anat. Rec.* **208**, 349–355 (1984).
 21. Nakajima, M., Irimura, T., Di, F. D., Di, F. N. & Nicolson, G. L. *Science* **220**, 611–613 (1983).
 22. Venkatasubramanian, K. & Solursh, M. *Dev. Biol.* **104**, 428–433 (1984).
 23. Ocalan, M., Goodman, S. L., Kuhl, U., Hauschka, S. D. & von der Mark, K. *Dev. Biol.* **125**, 158–167 (1988).
 24. Ontell, M. & Dunn, R. F. *Am. J. Anat.* **152**, 539–555 (1978).

ACKNOWLEDGEMENTS. We thank J. Price, C. Cepko and J.-F. Nicolas for retroviral producer cells and our colleagues for comments on the manuscript. This work was supported in part by a Lucille P. Markey Visiting Fellowship (S.M.H.) and the NIH and the Muscular Dystrophy Association of America (H.M.B.).

Two adjacent MyoD1-binding sites regulate expression of the acetylcholine receptor α -subunit gene

Jacques Piette, Jean-Louis Bessereau,
Monique Huchet & Jean-Pierre Changeux

UA Centre National de la Recherche Scientifique D1284, Neurobiologie Moléculaire, Département des Biotechnologies, Institut Pasteur, 75724 Paris Cédex 15, France

SEVERAL genes encoding putative myogenic regulatory factors have been isolated on the basis of their ability to convert non-muscle cells into myoblasts^{1–6}. Four of these genes code for nuclear proteins that belong to a larger family characterized by a conserved helix–loop–helix motif required for DNA-binding and dimerization⁷. At least one protein, MyoD1, can function as a transcription factor and activate muscle-specific genes during differentiation⁸. But the promoter of the δ -subunit gene of the mouse acetylcholine receptor (AChR) was recently reported to be functional in the absence of MyoD1 binding sites and it has been suggested that the genes coding for the AChR could be regulated independently of MyoD1 protein⁹. Here, we identify two functional MyoD1-binding sites in the muscle-specific enhancer of the chicken AChR α -subunit gene that are essential for full activity in transfected myotubes.

A minimal muscle-specific enhancer sequence of 36 base pairs (bp) has been defined in the chicken AChR α -subunit gene by transient expression experiments in various mouse myogenic cell lines^{10–12} and in primary cultures of chicken myotubes¹³. As two potential MyoD1-binding sites are present in this sequence, we used an affinity-purified glutathione S-transferase-MyoD1 fusion protein (Glu-MyoD)⁸ to probe enhancer frag-

ments in gel retardation experiments. No shift was observed with the bacterial enzyme alone or when a fragment containing an Sp1-factor binding site was probed with Glu-MyoD (results not shown). In contrast, two retarded bands were observed for the fragment ARIIpb containing the complete enhancer (Fig. 1a); the lower band is labelled I and the upper II (Fig. 2a, lane 1). Band I alone was detected with a fragment containing only the previously characterized element ARIIb of the enhancer¹³ (Fig. 2a, lane 4). Band II therefore probably represents a complex of two Glu-MyoD molecules, one bound to element ARIIb and a second to a more distal site.

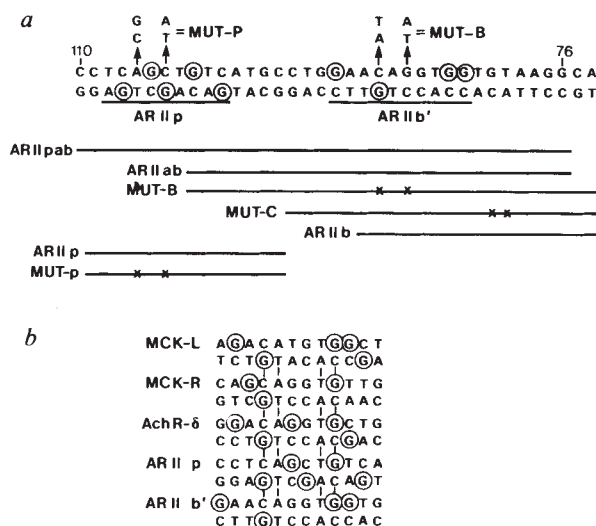
The two binding sites were characterized more precisely by means of methylation interference experiments. Guanines, the methylation of which interfered with the formation of complexes I and II (bands I and II, Fig. 2a), were analysed in Fig. 2d, and the results represented in Fig. 1a. We conclude that complex I corresponds mainly (if not exclusively) to Glu-MyoD protein bound to element ARIIb. Complex II corresponds effectively to the binding of two molecules, one bound to element ARIIb and a second to a more distal site overlapping a *PvuII* restriction site. We have called this latter site ARIIp and the downstream site ARIIb', because it is part of element ARIIb (ref. 13).

The two binding sites are compared with the previously characterized MCK-L, MCK-R and AChR- δ sites in Fig. 1b^{8,9}. Similarities exist in the nucleotide sequences and in the methylation interference patterns among the five sites. The close proximity of the two sites within the α -subunit gene enhancer is also remarkable. Similar motifs in both sites occur on the same side of the DNA helix, suggesting that direct interactions between two bound molecules may take place.

To evaluate the role of the two sites in α -subunit gene promoter activity, we constructed three mutants. The mutation of site ARIIp is called MUT-P, of site ARIIb' MUT-B and combination of both mutations MUT-BP (Fig. 1a). The effect of MUT-B and MUT-P on Glu-MyoD binding was tested in competition experiments. In both cases, a concentration of mutated fragment at least tenfold higher than wild-type fragment was necessary to displace the same amount of protein from the ARIIb' or ARIIp sites (Fig. 2b and 2c). We conclude that the point mutations decrease the affinity of Glu-MyoD for the α -subunit gene enhancer.

These mutations were then used to test the function of the previously characterized MyoD1 binding sites in transient expression experiments. 3T6 mouse fibroblasts were cotransfected with an expression vector for either MyoD1 or the related myogenin protein (plasmids, pEMSV-MyoD1 and pEMSV-myogenin, respectively)^{1,3} and a test vector containing either

FIG. 1 DNA sequence of the AChR α -subunit gene enhancer. a, The sequence represents nucleotides –110 to –74 relative to the transcription start site of the chicken AChR α -subunit gene¹⁰. Guanines whose methylation interfere with Glu-MyoD binding are circled, mutated base pairs in MUT-B and MUT-P are indicated with arrows. MUT-BP is the combination of both mutations. Double-stranded oligonucleotide fragments used in gel retardation and competition experiments are indicated below the sequence. Crosses represent mutated base pairs. A4-bp linker sequence was added 5' to facilitate subsequent cloning or 3'-end labelling of the oligonucleotide fragments. b, Comparison of MyoD1 binding sites in muscle-specific genes. Symbols are as in (a). Sequences for MCK-L and MCK-R are from ref. 8 and that for AChR- δ is from ref. 9. The conserved CA–TG motif is indicated.



wild-type or mutated promoter sequences of the AChR α -subunit gene (nucleotides -842 to +20 in p α AChCAT+; ref. 10) linked to the chloramphenicol acetyltransferase (CAT) gene (Fig. 3a). A minimal promoter containing only 45 bp upstream of the transcription start point was not transactivated under these conditions (not shown). In contrast, the p α AChCAT+ construct was transactivated by both MyoD1 and myogenin. Mutation of either site ARIIp or ARIIb' led to a significant inhibition of transactivation. Specifically, mutation of ARIIb' (MUT-B) had a stronger inhibitory effect than that of ARIIp (MUT-P). Mutation of both sites (MUT-BP) led to an almost complete loss of transactivation. These results show that at least two myogenic factors, MyoD1 and myogenin can transactivate the α -subunit gene promoter through the ARIIp and ARIIb' sites.

MyoD1 expression activated the α -subunit gene promoter as well as constructs containing several copies of the ARIIb element in front of the herpes simplex virus thymidine kinase (TK) promoter in several cell types: non-myogenic 3T6 mouse fibroblasts (Fig. 3a, and data not shown) or CV1 monkey epithelial cells (not shown) and myogenic C2.7 myoblasts (Fig. 3b). Figure 3b illustrates the dependence of the level of transactivation on the copy number of the MyoD1-target (compare ARIIb-3TKM containing three copies of ARIIb' and ARIIb-10TKM with 10 copies). Noteworthy is the fact that the presence of detectable levels of MyoD1 protein and messenger RNA in this cell line does not correlate with transcriptional activity^{14,15}. We speculate that expression of high levels of MyoD1 from the introduced vectors could either titrate out repressor molecules that interact with MyoD1, or might exhaust a repressing pathway, function-

ing for instance through covalent modifications of MyoD1.

We next investigated whether the two MyoD1 binding sites are critical for α -subunit gene promoter activity in the more physiological context of primary cultures of chick myotubes (Fig. 3c). Mutation of each MyoD1 site separately had only a moderately negative effect on promoter activity as determined by CAT-assay (~twofold). But the combination of both mutations produced a strong decrease in promoter activity (~20-fold). This effect was observed as early as one day after plating (results not shown), most probably perturbing the binding of factors appearing very early during *in vitro* differentiation of myoblasts. Thus, mutations of only four base pairs in sites ARIIb' and ARIIp strongly affect AChR α -subunit gene promoter function, highlighting the importance of the factors binding to these two sites in the tissue-specific expression of the promoter.

Data from (1) *in vitro* binding experiments using purified Glu-MyoD fusion protein, (2) transactivation experiments with expression vectors for MyoD1 and myogenin in several cell lines, and (3) the effect of specific point mutations of the MyoD1 binding sites on promoter activity, implicate myogenic factors of the MyoD1 family in the regulation of the AChR α -subunit promoter. Gel retardation and competition experiments with the ARIIp and ARIIb' sites suggest the presence in chicken myotubes of several MyoD1-like factors with closely related specificities (J.P., unpublished results). An attractive possibility is that among the different myogenic factors, some regulate 'constitutive' muscle genes such as muscle creatine kinase⁸ and others regulate 'electric activity-dependent' genes such as the AChR genes. This latter type of factor could form a

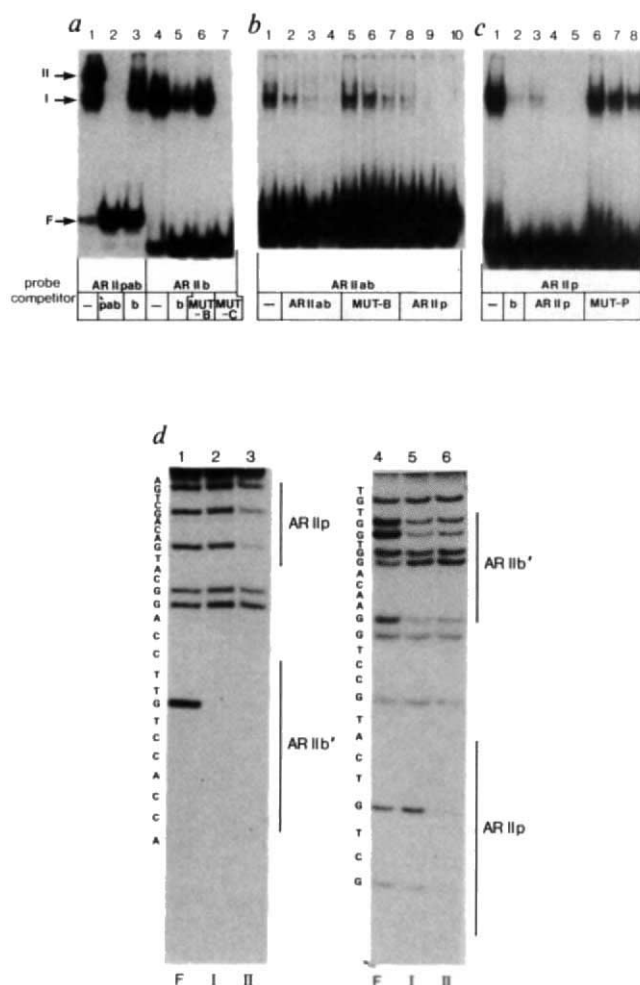


FIG. 2 MyoD1 binding to the AChR α -subunit gene enhancer. *a*, Probe ARIIpab (see Fig. 1) was used in lanes 1-3, probe ARIIb in lanes 4-7. A 200-fold molar excess of unlabelled fragment ARIIpab is used in lane 2, fragment ARIIb in lanes 3 and 5, fragment MUT-B in lanes 6 and fragment MUT-C in lane 7. One microgram affinity-purified Glu-MyoD protein was used in each experiment. *b*, Probe ARIIb (Fig. 1) was used in all lanes. A 20-fold, 100-fold and 200-fold molar excess was used for each competitor fragment: fragment ARIIb in lanes 2-4, fragment MUT-B in lanes 5-7 and fragment ARIIp in lanes 8-10; 0.4 μ g affinity purified Glu-MyoD protein was used in each experiment in the presence of 1 μ g poly(dI.dC). *c*, Probe ARIIp (Fig. 1) was used in all lanes. A 20-fold, 100-fold and 200-fold molar excess was used for each competitor fragment: fragment ARIIp in lanes 3-5 and fragment MUT-P in lanes 6-8. A 200-fold molar excess of fragment ARIIb was used in lane 2. 0.4 μ g of Glu-MyoD protein was used in each experiment. *d*, Methylation interference experiments. Fragment ARIIpab 5' labelled at the lower strand (Fig. 1a) was loaded in lanes 1-3, and at the upper strand in lanes 4-6. Free DNA (F, Fig. 2a) is loaded in lanes 1 and 4, DNA from complex I in lanes 2 and 5, and DNA from complex II in lanes 3 and 6. METHODS. *a-c*, Incubation and gel electrophoresis conditions were as described⁹. Oligonucleotide fragments were labelled by fill-in labelling with Klenow polymerase. *d*, Experiments were performed as described¹³, and the conditions for incubation and gel electrophoresis were as described for *a-c*. Fragment ARIIpab was used as probe. The methylated fragments were used in gel retardation experiments with 4 μ g Glu-MyoD in the presence of 1 μ g poly(dI.dC). The DNA was eluted from the retardation gels and subsequently cleaved at methylated purines with NaOH. DNA cleaved at methylated guanines was loaded onto a 15% denaturing polyacrylamide gel.

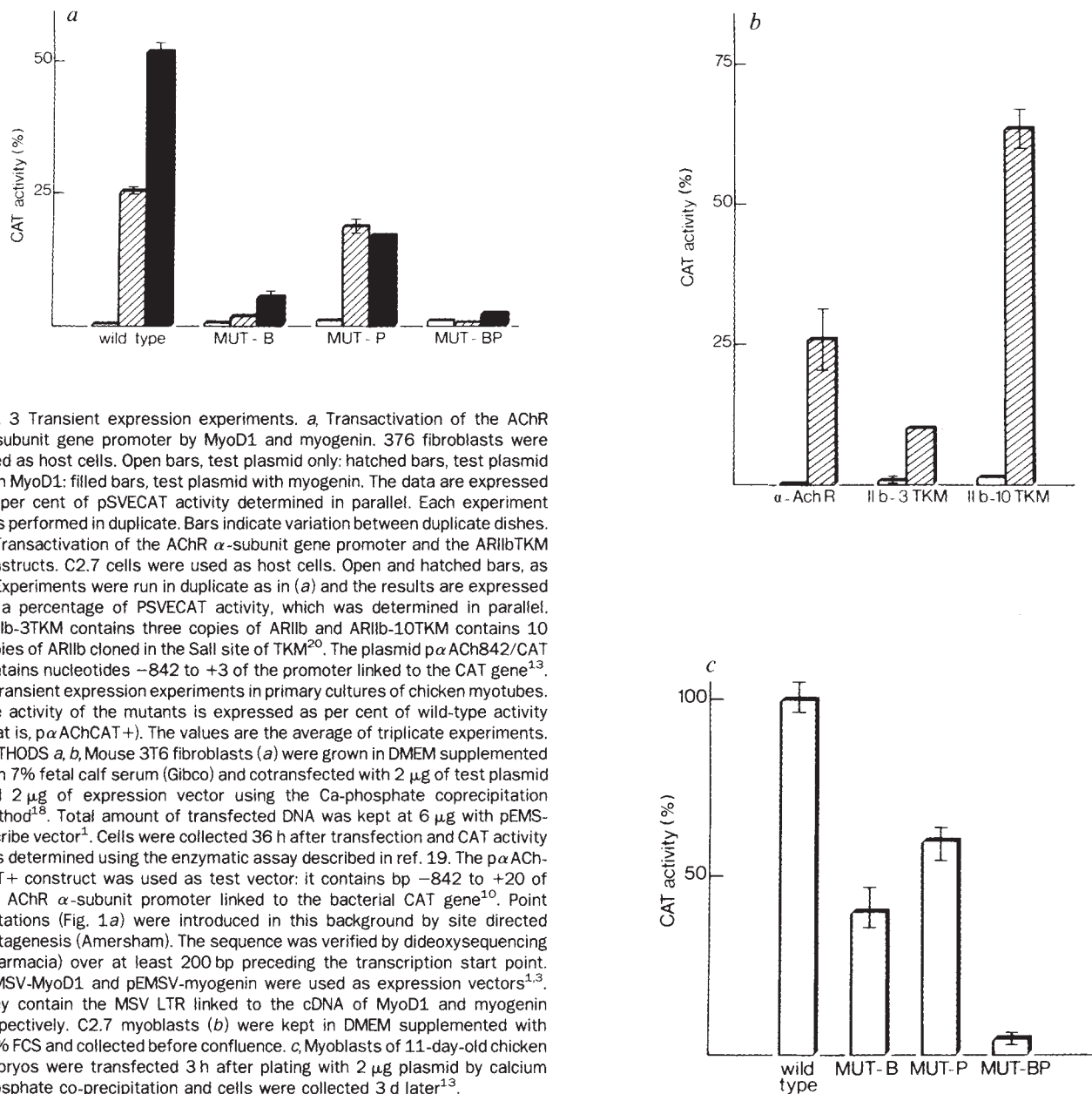


FIG. 3 Transient expression experiments. *a*, Transactivation of the AChR α -subunit gene promoter by MyoD1 and myogenin. 3T6 fibroblasts were used as host cells. Open bars, test plasmid only; hatched bars, test plasmid with MyoD1; filled bars, test plasmid with myogenin. The data are expressed as per cent of pSVECAT activity determined in parallel. Each experiment was performed in duplicate. Bars indicate variation between duplicate dishes. *b*, Transactivation of the AChR α -subunit gene promoter and the AR11bTKM constructs. C2.7 cells were used as host cells. Open and hatched bars, as *a*. Experiments were run in duplicate as in (*a*) and the results are expressed as a percentage of pSVECAT activity, which was determined in parallel. AR11b-3TKM contains three copies of AR11b and AR11b-10TKM contains 10 copies of AR11b cloned in the Sall site of TKM²⁰. The plasmid p α ACh842/CAT contains nucleotides -842 to +3 of the promoter linked to the CAT gene¹³. *c*, Transient expression experiments in primary cultures of chicken myotubes. The activity of the mutants is expressed as per cent of wild-type activity (that is, p α AChCAT+). The values are the average of triplicate experiments. **METHODS** *a*, *b*, Mouse 3T6 fibroblasts (*a*) were grown in DMEM supplemented with 7% fetal calf serum (Gibco) and cotransfected with 2 μ g of test plasmid and 2 μ g of expression vector using the Ca-phosphate coprecipitation method¹⁸. Total amount of transfected DNA was kept at 6 μ g with pEMSV-Scribe vector¹. Cells were collected 36 h after transfection and CAT activity was determined using the enzymatic assay described in ref. 19. The p α AChCAT+ construct was used as test vector: it contains bp -842 to +20 of the AChR α -subunit promoter linked to the bacterial CAT gene¹⁰. Point mutations (Fig. 1*a*) were introduced in this background by site directed mutagenesis (Amersham). The sequence was verified by dideoxysequencing (Pharmacia) over at least 200 bp preceding the transcription start point. pEMSV-MyoD1 and pEMSV-myogenin were used as expression vectors^{1,3}. They contain the MSV LTR linked to the cDNA of MyoD1 and myogenin respectively. C2.7 myoblasts (*b*) were kept in DMEM supplemented with 20% FCS and collected before confluence. *c*, Myoblasts of 11-day-old chicken embryos were transfected 3 h after plating with 2 μ g plasmid by calcium phosphate co-precipitation and cells were collected 3 d later¹³.

transcriptionally active complex with a labile protein whose synthesis is repressed by electrical activity or be itself subject to such regulation¹⁶. Characterization of the myogenic factors

that interact with the AChR genes should help to determine their role in differential expression of these genes in junctional and extra-junctional nuclei during endplate formation¹⁷. □

Received 30 January; accepted 29 March 1990.

1. Davis, R. L., Weintraub, H. & Lassar, A. B. *Cell* **51**, 987-1000 (1987).
2. Pinney, D. F., Pearson-White, S. M., Konieczny, S. F., Latham, K. E. & Emerson, C. P. Jr *Cell* **53**, 781-793 (1988).
3. Wright, W. E., Sassoon, D. A. & Lin, V. K. *Cell* **56**, 607-617 (1989).
4. Braun, T., Buschhausen-Denker, G., Bober, E., Tannich, E. & Arnold, H. H. *EMBO J.* **8**, 701-709 (1989).
5. Rhodes, S. J. & Konieczny, S. F. *Genes Dev.* **3**, 2050-2061 (1989).
6. Colmenares, C. & Stavnezer, E. *Cell* **59**, 293-303 (1989).
7. Murre, C., McCaw, P. S. & Baltimore, D. *Cell* **56**, 777-783 (1989).
8. Lassar, A. B. *et al.* *Cell* **58**, 823-831 (1989).
9. Baldwin, T. J. & Burden, S. J. *Nature* **341**, 716-720 (1989); correction *Nature* **345**, 364 (1990).
10. Klarsfeld, A., Daubas, P., Bourachot, B. & Changeux, J.-P. *Molec. cell. Biol.* **7**, 951-955 (1987).
11. Daubas, P. *et al.* *Nucleic Acids Res.* **16**, 1251-1271 (1988).
12. Wang, Y., Xu, H. P., Wang, X. M., Ballivet, M. & Schmidt, J. *Neuron* **1**, 527-534 (1988).

13. Piette, J., Klarsfeld, A. & Changeux, J.-P. *EMBO J.* **8**, 687-694 (1989).
14. Tapscott, S. J. *et al.* *Science* **242**, 405-441 (1988).
15. Montarras, D., Pinset, C., Chelly, J., Kahn, A. & Gros, F. *EMBO J.* **8**, 2203-2207 (1989).
16. Duclert, A., Piette, J. & Changeux, J.-P. *Proc. natn. Acad. Sci. U.S.A.* **87**, 1392-1395 (1990).
17. Changeux, J. P. *et al.* *NATO ASI Series Vol. H32* (Springer, Berlin, 1989).
18. Graham, F. L. & Vander Eb, A. J. *Virology* **52**, 456-467 (1973).
19. Seed, B. & Sheen, J.-Y. *Gene* **67**, 271-277 (1988).
20. Tronche, F. *et al.* *Molec. Biol. Med.* (in the press).

ACKNOWLEDGEMENTS. We thank Melissa Lakich and Bob Whalen for Glu-MyoD protein; Andrew Lassar and Woody Wright for plasmids pEMSV-MyoD1 and pEMSV-myogenin; André Klarsfeld for help with the constructs; Didier Montarras and Christian Pinset for discussion; Moshe Yaniv, Aymeric Duclert, Joe Hill and Ralph Lauffer for critically reading the manuscript; and Chantal Le Poupon and Henriette Nuret for technical assistance. J. B. is a pre-doctoral fellow of the ARC. This study was supported by the CNRS, the Collège de France, INSERM and the DRET.

Genetic organization of a chimpanzee lentivirus related to HIV-1

Thierry Huet, Rémi Cheyrier, Andreas Meyerhans, Georges Roelants* & Simon Wain-Hobson

Laboratoire de Biologie et Immunologie Moléculaires des Rétrovirus, Institut Pasteur, 28 Rue de Dr Roux, 75724 Paris cedex 15, France

* Centre International de Recherches Médicales de Franceville, BP 769, Franceville, Gabon

SIMIAN immunodeficiency viruses have been isolated from four species of monkey, the 'captive' macaque¹ and mangabey²⁻⁴ and the 'feral' African green monkey⁵ and mandrill⁶. While none of these viruses is a replica of HIV-1, the macaque⁷ and mangabey⁸ viruses represent correct genetic models for HIV-2, possessing exactly the same complement of genes. Recently a lentivirus has been identified in two wild chimpanzees (*Pan troglodytes troglodytes*) in Gabon, west equatorial Africa, and isolated from one of them⁹. This virus is referred to as SIV_{CPZ}. Sera from these animals cross reacted with all the HIV-1 proteins including the envelope glycoproteins. Here, we describe the molecular cloning and sequencing of an infectious proviral clone of SIV_{CPZ}. The overall genetic organization was the same as that of HIV-1, but phylogenetic analysis revealed that the sequence was more divergent than any HIV-1 sequence reported so far. The *vpu* gene product, found only in the type 1 viruses, was particularly different (64% divergent to HIV-1_{BRU}) suggesting that the SIV_{CPZ} represents a distinct subtype. These findings indicate that there is a larger pool of simian lentiviruses than previously suspected and revives debate as to the origins of HIV-1.

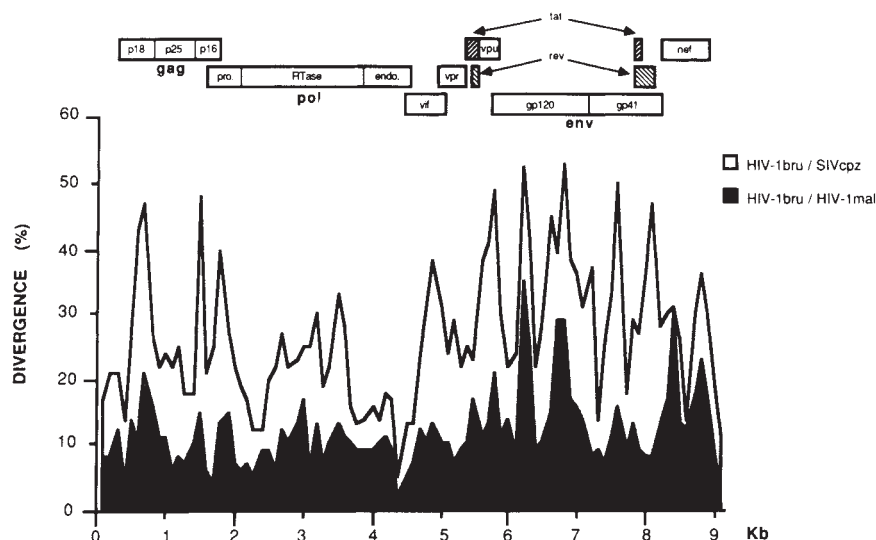
A phage λ genomic library was constructed from total DNA of SIV_{CPZ}-infected human lymphocytes. On screening with a HIV-1 probe, three recombinants (1a, 30 and 46a) carrying a complete SIV_{CPZ} provirus were isolated. The three clones had identical *Sma*I, *Xba*I, and *Bam*HI restriction profiles as well as eight *Hind*III restriction sites, one of which was lacking in clone 46a. Fragments containing the complete provirus from clones 1a and 46a were subcloned into plasmid pKP55 using unique *Eco*RI restriction sites in the cellular flanking sequences. On transfection into the human colon carcinoma cell line SW480, both clones were able to produce virus, with reverse transcriptase titres peaking at 7 and 14 days respectively for clone 1a and 46a (data not shown). Clone 1a virus proved to be infectious either by co-cultivation of the transfected SW480 cells with human peripheral blood mononuclear cells (PBMCs) or by passage of cell-free supernatant to donor PBMCs. Because of this and its more rapid kinetics, clone 1a was selected for sequencing (see Fig. 1 legend, ref. 10).

The SIV_{CPZ} genome proved to be more closely related to HIV-1 group than to any of the simian lentiviruses. Thus the organization of the SIV_{CPZ} genome was 5' *gag-pol-vif-vpr-tat-rev-vpu-env-nef* 3'. The presence of a HIV-1-specific *vpu* gene and the absence of a *vpx* gene, common to HIV-2 and most SIVs, clearly show that SIV_{CPZ} is the first non-human lentivirus to have the same genetic structure as HIV-1. Table 1 gives the percentage amino acid sequence similarities for pairwise comparisons of all gene products of SIV_{CPZ} to those from other human and simian lentiviruses. The values confirm that quantitatively SIV_{CPZ} has, as nearest relative, the HIV-1 virus group. Nonetheless, the SIV_{CPZ} genome is more divergent than any of the HIV-1 isolates, including African and European strains, and yet is also equidistant to all of these.

To better appreciate the degree of divergence of the SIV_{CPZ} genome, a pairwise comparison of the SIV_{CPZ}/HIV-1_{BRU}¹¹

FIG. 1 Organization of the aligned SIV_{CPZ} genome and sequence divergence with respect to HIV-1_{BRU}. The HIV-1_{BRU}, HIV-1_{MAL} and SIV_{CPZ} nucleotide sequences were aligned. All gaps and corresponding sequences were eliminated. The percentage sequence variation was calculated for the HIV-1_{BRU}/HIV-1_{MAL} (dark shaded area in graph) and HIV-1_{BRU}/SIV_{CPZ} (lightly shaded area in graph) pairs over a window of 100 bases. The windows did not overlap. Gaps apart the BRU/MAL and BRU/CPZ genomes are 11.8% and 26.9% divergent respectively. In every case the SIV_{CPZ} sequence was always the more divergent of the two. Some parts of the *gag* gene, notably p18 and p16 are almost as divergent as parts of *env* at the nucleic acid level.

METHODS. Total DNA was partially digested with *Bgl*II. Fragments of 10–20 kilobases (kb) were selected from a sucrose density gradient, purified and ligated into *Bam*HI-cleaved λ EMBL3 DNA. After packaging *in vitro*, a phage library of 1.5×10^6 clones was plated out on *Escherichia coli* P2 392. Plaques were screened *in situ* with a complete probe of HIV-1_{OVI}²². Three clones carrying full length SIV_{CPZ} proviruses were studied in detail. Clones 1a and 46 were subcloned into the pKP55 vector (supplied by Keith Peden). Plasmid DNA (10 μ g) was transfected onto 2×10^6 SW480 cells by a modified CaPO_4 method²⁵. Near the peak of reverse transcriptase, supernatant was taken, clarified and used to infect PBMCs. A third of the SW480 cells were irradiated and co-cultivated with PBMCs. Reverse transcriptase in the culture supernatants confirmed that clone 1a represented an infectious molecular clone. Clone 1a DNA was sonicated as described¹¹ and fragments of 600–1000 base pairs (bp) isolated and cloned into the dephosphorylated *Sma*I site of M13mp18 RF DNA. Templates were sequenced using Taq polymerase and fluorescent primers and the products resolved using an Applied Biosystems 370A DNA sequencer¹⁰. The sequence was compiled as described²⁶. Two regions were poorly represented in the contig. Accordingly, oligonucleotide



primers were synthesized allowing amplification of regions 2084–2273 and 6215–7132. These regions were amplified by the polymerase chain reaction²⁷ (PCR) from 100 ng of plasmid 1a DNA in 10 cycles. Amplified DNA was cleaved by *Bam*HI (this restriction site having been added to the PCR primers) and cloned into the *Bam*HI site of M13mp18 RF DNA and sequenced. For each region two or more clones were sequenced, ruling out Taq polymerase error. Each base in the SIV_{CPZ} sequence was sequenced on average 8.9 times. The flanking cellular DNA represented non-coding sequences. Like HIV-1, a 5-bp direct repeat of cellular DNA was formed on integration. The SIV_{CPZ} provirus was 9,811 bp long with the following base composition: T 22.7%; C 18.7%; G 24.5%; A 34.2%. This sequence has been filed with the EMBL database under accession number X52154.

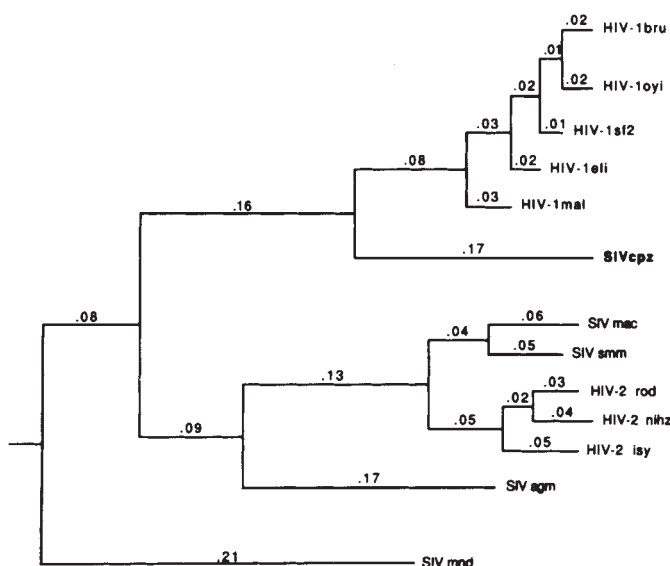


FIG. 3 Minimum length evolutionary tree for the gag region showing the relationship of SIV_{CPZ} to the known HIV and SIV lentiviruses. The total number of sites examined was 1,473 of which 967 were variable. The consistency index was 0.675. The lengths of the horizontal lines are proportional to the minimum number of single nucleotide substitutions required to generate the observed variation (also indicated by the numeric branch lengths above each line). The additional sequences not referred to in Table 1 are HIV-1_{SF2}²⁹, SIV_{SMM} (sooty mangabey)⁸, HIV-2_{NIHZ}²⁹ and HIV-2_{ISY}²⁹. METHODS. The protein sequences were aligned using a clustering and pattern analysis algorithm from the amino-acid class covering utility (AACC)¹⁹. Corresponding nucleotide sequence alignments were derived from the protein alignments and subjected to multiple parsimony using the PAUP program²⁰. The bootstrap analysis used the DNABOOT program in the PHYLIP 3.21 package³⁰.

The transmembrane protein (TM) is nearly as divergent as the SU (Table 1). This is unusual for HIV-1 but not for HIV-2. The SIV_{CPZ} *rev* protein is also quite divergent (Fig. 1 and Table 1). As most of the *rev* coding region overlaps with the C-terminal half of TM, this could explain why the TM sequence is so variable. Finally, in every HIV-1 sequence to date, the *env* stop codon (TAA) forms part of the optimal *nef* start sequence (TAAGATGG in which the 'Kozak' motif is underlined). For SIV_{CPZ} an additional 12 bases separate the two open reading frames.

As anticipated, the envelope variable regions were indeed hypervariable. Apart from the cysteine and a few flanking residues, as well as the conserved V3 GPG tripeptide, it seems that virtually any hydrophilic amino acid sequences can be accommodated. If so, then why do isolates from AIDS patients still show sequence similarities in these regions? A simple explanation could be that the global AIDS epidemic is derived from one particular strain, reflected by the HIV-1_{BRU-MAL} family, the differences among them representing only 15–20 years of divergence. In a similar manner, many of the North American strains are closely related to strain HIV-1_{MN}¹⁸, perhaps because this, or a related strain, was one of the first to enter the United States.

A phylogenetic tree analysis^{19,20} was undertaken of the SIV_{CPZ} nucleotide sequence. The minimal tree for the *gag* region is given in Fig. 3. The topology of the tree was the same for all regions studied. Moreover, bootstrap analysis of the *nef* tree (not shown) confirmed this position for the SIV_{CPZ} sequence in 50 out of 50 trials. It is clear that SIV_{CPZ} branched before any of the HIV-1 strains sequenced so far. The chimpanzee and HIV-1 sequences were more divergent than SIV_{MAC} or SIV_{SMM} were to HIV-2 (Fig. 3). In view of the particularity of SIV_{CPZ}, its original designation⁹, it might be more informative to refer to this virus as chimpanzee immunodeficiency virus, CIV.

In Gabon, only 2 out of 83 chimpanzees tested were seropositive, indicating that SIV_{CPZ} is not widely dispersed in this region. Given their troop behaviour and fear of water, it is possible that SIV_{CPZ} infection may be highly restricted geographically. Of more than 250 chimpanzees caught over the last 15–20 years in West Africa, none was seropositive. This might help explain the absence of naturally infected chimpanzees in captivity in the United States as virtually all are of West African origin (A. Prince, personal communication).

That the two chimpanzees were both infants and in relatively good health should not be surprising, for if they suffered from

anything resembling AIDS or AIDS-related complex, they would not have long survived predators in the jungle. Intriguingly none of the 'feral' SIVs, the African green monkey and mandrill viruses, is associated with disease, perhaps for the same reason. Ironically simian AIDS has only been described among captive monkeys where living conditions and general hygiene are much better. Consequently, discussion of the relative pathogenicity of SIV_{CPZ} is not prudent. It is, however, an intriguing question as there is relatively little AIDS in this region of Africa.

Inevitably the similarity between the HIV-1 and SIV_{CPZ} genomes elicits speculation as to their evolutionary relationship and the origins of the current AIDS epidemic. Obviously there are HIV-1-like viruses in chimpanzees. It is not possible to conclude that SIV_{CPZ} was the precursor to HIV-1, if indeed infection ever passed in that direction. Even given this premise, the *vpu* data indicate that SIV_{CPZ} was not the immediate precursor. Clearly more data on SIV_{CPZ} prevalence and genetics would be useful.

In only a few years, lentiviruses have been isolated from African green monkeys, mandrills and now chimpanzees. It may be only a matter of time before other HIV-1-like lentiviruses will be isolated. An extensive serological study of simians, perhaps using the HIV-1 *vpu* sequence might yield just such a virus. □

Received 12 February, accepted 22 March 1990.

1. Daniel, M. D. *et al. Science* **228**, 1201–1204 (1985).
2. Lowenstein, L. J. *et al. Int. J. Cancer* **38**, 563–574 (1986).
3. Murphy-Corb, M. *et al. Nature* **321**, 435–437 (1986).
4. Fultz, P. N. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 5286–5290 (1986).
5. Ohta, Y. *et al. Int. J. Cancer* **41**, 115–122 (1988).
6. Tsujimoto, H. *et al. J. Virol.* **62**, 4044–4050 (1989).
7. Chakrabarti, L. *et al. Nature* **328**, 543–547 (1987).
8. Hirsch, V. M., Olmsted, R. A., Murphy-Corb, M., Purcell, R. H. & Johnson, P. R. *Nature* **339**, 389–392 (1989).
9. Peeters, M. *et al. AIDS* **3**, 625–630 (1989).
10. Connell, C. *et al. Bio Techniques* **5**, 342–348 (1987).
11. Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S. & Mizon, M. *Cell* **40**, 9–17 (1985).
12. Allison, M., Wain-Hobson, S., Montagnier, L. & Sonigo, P. *Cell* **46**, 63–74 (1986).
13. Guyader, M. *et al. Nature* **326**, 662–669 (1987).
14. Emerman, M., Guyader, M., Montagnier, L., Baltimore, D. & Muesing, M. *EMBO J.* **6**, 3755–3760 (1987).
15. Malim, M. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 8222–8226 (1989).
16. Klimkait, T., Strebel, K., Hoggan, M. D., Martin, M. A. & Orenstein, J. M. *J. Virol.* **64**, 621–629 (1990).
17. Hammerschmidt, W., Sugden, B. & Baichwal, V. R. *J. Virol.* **63**, 2469–2475 (1989).
18. Gurgo, C. *et al. Virology* **164**, 531–536 (1988).
19. Smith, R. F. & Smith, T. F. *Proc. natn. Acad. Sci. U.S.A.* **87**, 118–122 (1990).
20. Smith, T. F., Srinivasan, A., Schochetman, G., Marcus, M. & Myers, G. *Nature* **333**, 573–575 (1988).
21. Wilbur, W. J. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 726–730 (1983).
22. Huet, T., Dazza, M. C., Brun-Vézinet, F., Roelants, G. E. & Wain-Hobson, S. *AIDS* **3**, 707–715 (1989).

23. Fukasawa, M. *et al. Nature* **333**, 457–461 (1988).
24. Tsujimoto, H. *et al. Nature* **341**, 539–541 (1989).
25. Chen, C. & Okayama, H. *Molec. cell. Biol.* **7**, 2745–2752 (1987).
26. Staden, R. *Nucleic Acids Res.* **10**, 4731–4751 (1982).
27. Saiki, R. K. *et al. Science* **239**, 487–491 (1988).
28. Kyte, J. & Doolittle, R. *J. molec. Biol.* **157**, 105–132 (1982).
29. Myers, G. *et al. Human Retroviruses and AIDS* (Los Alamos National Laboratory, New Mexico, 1989).
30. Felsenstein, J. A. *Rev. Genet.* **22**, 521–565 (1988).

ACKNOWLEDGEMENTS. We are indebted to Drs Gerald Myers, Kersti Macinnes and Jeffrey Lawrence for the phylogenetic analysis and comments on the manuscript. We thank Drs Martine Peeters and Eric Delaporte for SIV_{CR2} and Dr Michel Dubos (D.R.E.T) for support. T.H. was funded by the Institut Mérieux; R.C. from the Fondation pour la Recherche Médicale and A.M. from the Centre International des Etudiants et Stagiaires.

How different eukaryotic transcriptional activators can cooperate promiscuously

Young-Sun Lin, Michael Carey, Mark Ptashne & Michael R. Green

Department of Biochemistry and Molecular Biology, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

A STRIKING characteristic of many different eukaryotic transcriptional activators is their ability to activate gene expression synergistically. Thus, for example, the rat glucocorticoid receptor and the yeast activator GAL4 cooperatively activate transcription of a mammalian gene bearing binding sites for each of the proteins¹: activation by both activators is greater than the sum of the effects of each working alone. It would seem unlikely that these two proteins from such different organisms directly interact; rather, the idea has been suggested that these and at least some other eukaryotic activators can work synergistically by simultaneously touching some part of the transcriptional machinery². An important prediction of this idea is that synergy between two such activators would be seen under conditions where each is present at concentrations sufficiently high to saturate its site on DNA. In this paper we use transcription *in vitro* to confirm that prediction using a derivative of the yeast activator GAL4 and the mammalian transcription factor ATF. The accompanying paper³ describes a similar conclusion comparing the effects of singly and multiply bound GAL4 molecules.

We have previously demonstrated that a GAL4 derivative synthesized in and purified from *Escherichia coli* can function synergistically with the mammalian activator ATF in a HeLa cell nuclear extract⁴. Here we study this phenomenon further and show that synergism is observed under conditions in which the binding sites for ATF and the GAL4 derivative are saturated.

The template used in these transcription experiments contained a single ATF site immediately adjacent to the TATA box of the adenovirus E4 gene; five GAL4 sites, termed 17-mers, were positioned 160 base pairs (bp) upstream (see Fig. 1). Transcription was carried out in HeLa cell nuclear extracts depleted of endogenous ATF by sequence-specific DNA affinity chromatography^{5,6}. The *E. coli*-derived GAL4 derivative GAL4(1–147) + AH (ref. 7) and purified mammalian ATF were added to the extracts as indicated and transcription was measured by primer extension⁴. Binding of GAL4(1–147) + AH and ATF in the extracts was examined by DNase I footprinting using a ³²P-labelled template DNA fragment.

Increasing amounts of purified ATF or GAL4 (1–147) + AH were added to an ATF-depleted extract and transcriptional activity was determined in the presence of an excess of the other activator. The autoradiograph of such a transcription experiment is shown in Fig. 1 and the data quantitated and plotted in Fig. 2. Figure 2a shows the transcription elicited by various amounts of GAL4(1–147) + AH in the presence and absence of excess ATF. This figure also shows the occupancy of the GAL4 sites as measured by DNase I footprinting. Figure 2b shows the transcription elicited by ATF in the presence and absence of excess GAL4(1–147) + AH. This figure also shows the occupancy of the ATF site as measured by DNase I footprinting.

Two results are evident from the experiment of Fig. 2. First, in each case, activation paralleled site filling as assayed by DNase I footprinting. Second, taking Fig. 2a and b together, addition of an excess of one activator resulted in a synergistic increase in transcription at all concentrations of the other activator. In particular, as shown by the bar graph of Fig. 3, this synergy was observed when both activators were present at concentrations sufficient to saturate their sites on DNA. In Fig. 3 the hatched bar represents transcription at the highest concentration of ATF tested in the absence of GAL4(1–147) + AH; the grey bar represents transcription at the highest concentration of GAL4(1–147) + AH tested in the absence of ATF; the black bar represents transcription at the highest concentrations of both activators. The white bar represents the expected value if the response to the two activators were additive. Because the black bar is greater than the white bar we say that ATF and GAL4(1–147) + AH are working synergistically.

As purified ATF comprises multiple DNA-binding species⁶, we cannot assess the relative contribution of each polypeptide to the observed transcriptional activity. But this does not alter our conclusions, which are based on the observation that the GAL4 and ATF binding sites are saturated.

Our experiments describe synergistic activation under conditions that render cooperative DNA binding of the activators irrelevant, that is, at a concentration of each activator sufficient to fill its sites. We have also obtained similar results by adding

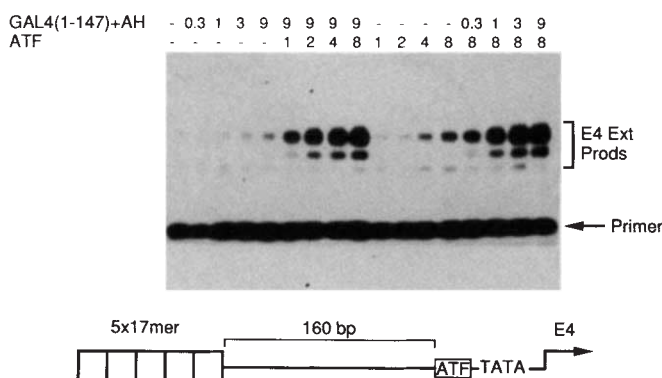
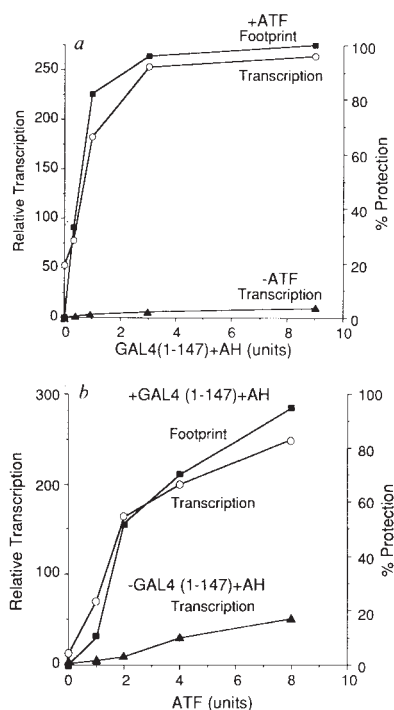


FIG. 1 Titration of ATF and GAL4 in an ATF-depleted nuclear extract. Transcription mixtures contained ATF-depleted nuclear extracts and the indicated concentrations (DNase I footprinting units) of either ATF or GAL4(1–147) + AH. Transcription products were measured by primer extension. The primer-extension products were fractionated on a denaturing polyacrylamide gel and made visible by autoradiography. The DNA template used in the experiment is shown schematically below the autoradiograph.

METHODS. The preparation of ATF oligonucleotide column and purification of ATF were as previously described^{5,6}. Nuclear extract (2 ml) was adjusted to 0.3 M KCl, passed through 1 ml of the ATF resin pre-equilibrated with buffer D (ref. 15) containing 0.3 M KCl. The flow-through was collected and reloaded onto the ATF oligonucleotide resin. The second flow-through was collected as the ATF-depleted nuclear-extract and dialysed against buffer D containing 0.05 M KCl. The ATF-depleted nuclear extracts contain no detectable ATF activity as measured by DNase I footprinting and transcription assays; these extracts could support normal levels of transcription from the adenovirus major late promoter, indicating that they contained all general transcription factors (data not shown). All other methods were as previously described⁴.



GAL4(1-147) + AH to a crude HeLa cell nuclear extract in which endogenous ATF is present at a saturating concentration (ref. 4, and data not shown). We imagine that this synergy is a consequence of interaction of each of the bound activators with some component or components of the transcriptional machinery rather than with each other. We might further consider two alternatives: the activators might simultaneously contact the same target as illustrated in Fig. 4, or they might contact different targets and affect different steps in the activation process. The accompanying paper describes synergistic activation by multiply-bound copies of a single activator, a result suggesting multiple contacts with a single target or with multiple copies

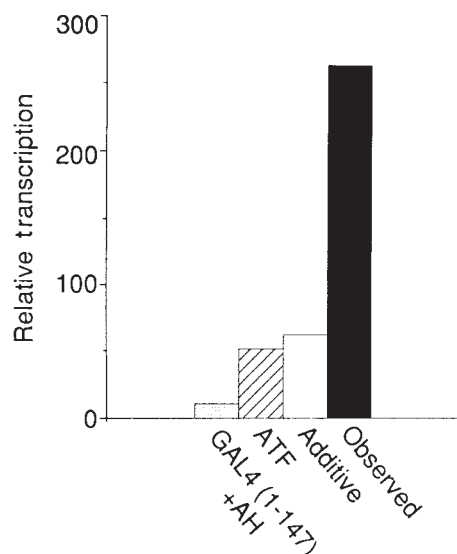


FIG. 3 Bar graph of synergism. The autoradiograph shown in Fig. 1 was scanned with a densitometer and the relative areas plotted as a bar graph. The y axis is relative transcription; the amount of transcription stimulated by the highest concentrations of ATF and GAL4(1-147) + AH are indicated; additive indicates the sum of the transcription signals elicited by GAL4(1-147) + AH and ATF assayed separately; observed indicates the amount of transcription actually measured when the two activators were assayed in the same reaction mixture.

FIG. 2 Graphical comparison of DNase I footprinting (■) and transcription (○, ▲). a, GAL4 titration plus or minus saturating ATF. The autoradiograph shown in Fig. 1 was scanned with a densitometer and the normalized areas for each point plotted. The Y axis indicates relative transcription and per cent protection assayed in parallel by DNase I footprinting; a value of one has been arbitrarily assigned to the level of transcription seen in the absence of either activator. The x axis indicates the amount of GAL4(1-147) + AH in DNase I footprinting units. b, ATF titration plus or minus saturating GAL4(1-147) + AH. The figure is organized as in a except that titration was with ATF.

METHODS. Autoradiographs were scanned with a Helena densitometer and quantitated. For the transcription data (Fig. 1) dilutions of the primer were used to create a standard curve. For the DNase I footprints (not shown) the ratio of the protected bands to the unaffected bands was determined at each point in the titration. This ratio was divided by that determined in a control reaction lacking proteins. Plotted is $(1 - \text{the quotient}) \times 100$, the per cent of DNA protected by the protein.

of a single target. Previous reports^{8,9} suggest that both of the activators used in these experiments interact with the transcription factor, TFIID.

We note that, should the activators simultaneously contact a third component, they could bind DNA cooperatively, a phenomenon that could be relevant at low activator concentration. Cooperative DNA binding has a crucial role in the action of λ repressor¹⁰, for example, but in that case the cooperativity results from direct repressor-repressor interaction¹¹.

Many eukaryotic promoters contain multiple binding sites for various activators, the combinations of which vary amongst different promoters. In several cases it has been shown that mutation of one or more of these binding sites can be compensated for by duplicating one or more of the remaining sites

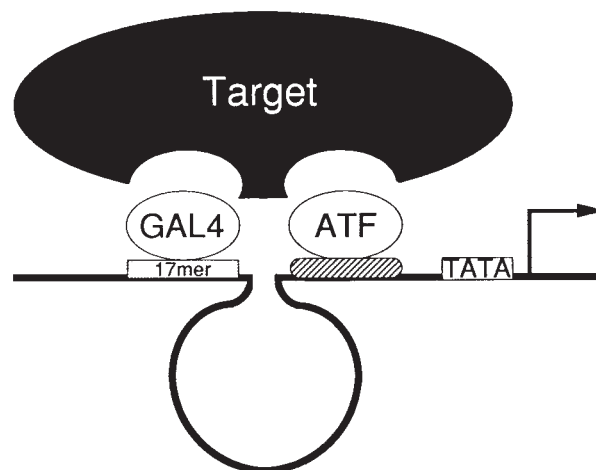


FIG. 4 Schematic representation of heterologous activators simultaneously contacting a common target. The diagram shows molecules of GAL4(1-147) + AH and ATF, bound to their sites on DNA, simultaneously contacting an as-yet-unidentified target protein; the intervening DNA loops out to accommodate the reaction. Our experiments indicate that as many as five activators may simultaneously interact with the target.

(see, for example, refs 12–14). We imagine that at least some of the activators that bind to these sites in eukaryotic promoters work synergistically by the mechanism described in this paper. □

Received 17 November 1989; accepted 26 March 1990.

1. Kakidani, H. & Ptashne, M. *Cell* **52**, 161–167 (1988).
2. Ptashne, M. *Nature* **335**, 683–689 (1989).
3. Carey, M. F., Lin, Y.-S., Green, M. R. & Ptashne, M. *Nature* **345**, 361–364 (1990).
4. Lin, Y.-S., Carey, M., Ptashne, M. & Green, M. R. *Cell* **54**, 659–664 (1988).
5. Lin, Y.-S. & Green, M. R. *Proc. natn. Acad. Sci. U.S.A.* **86**, 109–113 (1989).
6. Hai, T. Y., Liu, F., Allegretto, E. A., Karin, M. & Green, M. R. *Genes Dev.* **2**, 1216–1226 (1988).
7. Giniger, E. & Ptashne, M. *Nature* **330**, 670–672 (1987).
8. Horikoshi, M., Carey, M., Kakidani, H. & Roeder, R. G. *Cell* **54**, 665–669 (1988).
9. Horikoshi, M., Hai, T., Lin, Y.-S., Green, M. R. & Roeder, R. G. *Cell* **54**, 1033–1042 (1988).
10. Johnson, A. D., Pabo, C. O. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5061–5065 (1979).
11. Hochschild, A. & Ptashne, M. *Cell* **44**, 681–687 (1986).
12. Ondek, B., Gloss, L. & Herr, W. *Nature* **333**, 40–45 (1988).
13. Fromenthal, C., Kanno, M., Nomiya, H. & Chambon, P. *Cell* **54**, 943–953 (1988).
14. Schaffner, G., Schirm, S., Muller-Baden, B., Weber, F. & Schaffner, W. *J. molec. Biol.* **201**, 81–90 (1988).
15. Dignam, J. D., Martin, P. L., Shastry, B. S. & Roeder, R. G. *Meth. Enzym.* **101**, 582–598 (1983).

ACKNOWLEDGEMENTS. This work was supported by grants from the National Institutes of Health and American Cancer Society to M.P. and M.R.G. M.C. was supported by a postdoctoral fellowship from the American Cancer Society.

A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives

Michael Carey, Young-Sun Lin, Michael R. Green & Mark Ptashne

Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

IN prokaryotes and eukaryotes many gene activators work synergistically. For example, two dimers of λ repressor interact to promote binding of these proteins to DNA, a reaction that is crucial at the repressor concentrations found in lysogens¹. In this case one of the bound dimers activates transcription, evidently by touching RNA polymerase². In another example, the yeast transcriptional activator GAL4, which can stimulate transcription in many eukaryotes, binds to multiple sites on DNA to activate transcription synergistically; the presence of two such sites can elicit a level of transcription more than twice that found with a single site³. In this paper we show that synergistic activation by each of several GAL4 derivatives involves a mechanism different from that illustrated by the λ repressor: multiple activator molecules can work synergistically under conditions in which their binding sites on DNA are saturated. The accompanying paper shows that under similar conditions of activator excess, GAL4 derivatives work synergistically with a heterologous mammalian gene activator⁴. These results support the idea that eukaryotic activators can cooperate not by directly interacting but by simultaneously touching some component(s) of the transcriptional machinery.

Figure 1 compares activation *in vivo* using GAL4(1–147)+II (ref. 5) and GAL(1–147)+VP16 (ref. 6), a weak and a strong activator, respectively. These GAL4 derivatives bear identical DNA-binding regions, but differ in their activation regions: II is derived from GAL4 and VP16 from the herpes viral protein of that name. Each template contains the TATA region from the adenovirus E1b promoter fused to the reporter gene for chloramphenicolacetyltransferase CAT (ref. 7) and bears either 0, 1, 2, 5 or 11 17-base pair (bp) GAL4 binding sites⁸ upstream from the TATA box. These templates were co-transfected into Chinese hamster ovary (CHO) cells with an effector plasmid directing the synthesis of one or the other activator, and CAT

activity was taken as a measure of transcription. Figure 1a shows CAT assays obtained in such an experiment, and the tabulation in Fig. 1b compares the response of pairs of templates bearing different numbers of GAL4 DNA binding sites. For each pair in Fig. 1b, the black bar represents transcription from the template bearing the smaller number of sites and the hatched bar that from the template bearing the larger number of sites; the white bar represents in each case the expected value if the response to the multiple bound activators were additive with respect to the template with the smaller number of sites. Thus the ratio of the heights of the cross-hatched to the white bars gives a measure of synergism. Using GAL4(1–147)+VP16, for example, the response to five sites is at least 3.5 times greater than five times the response of a template bearing a single site. Using GAL4(1–147)+II, the response of a template bearing five sites is at least 33-fold greater than 2.5 times the response of a template bearing two sites. We take these results to mean that both activators are working synergistically (cooperatively). Two additional points to note are: first, the differences in the responses to the strong and weak activators diminish as the number of sites increases. Thus for example, on a template bearing only two sites the difference is at least 100-fold, whereas on a template bearing eleven sites the difference is only about twofold. Second, for both the strong and the weak activators, as the number of sites is increased, a point is reached at which synergy is no longer observed; this point requires more sites for the weak activator (5–11) than it does for the strong activator (2–5).

The results of the experiment described in Fig. 1c argue that the experiments of Fig. 1a were carried out at conditions in which the activator-binding sites were saturated. Figure 1c shows a typical titration curve generated by co-transfecting cells with a constant amount of reporter plasmid and increasing amounts of effector plasmid. As the effector concentration increased, transcription rose and reached a plateau at between 5 and 10 μ g of effector plasmid added. The experiments of Fig. 1a and b were all carried out using 10 μ g of effector plasmid, an amount sufficient for maximal transcription.

The experiments of Figs 2 and 3, carried out *in vitro*, further support the idea that synergy is observed under conditions in which the activators saturate their DNA-binding sites. For these experiments we used GAL4 derivatives synthesized in and purified from *Escherichia coli*. In addition to GAL4(1–147)+VP16 we used two weak activators: GAL4(1–147)+AH (ref. 9) bears a short peptide designed to form an amphipathic α -helix and GAL4(1–147) bears a cryptic activating region that weakly activates transcription in mammalian cell extracts¹⁰ but not in mammalian and wild-type yeast cells. Each template bears 0, 1, 2, 5 or 9 GAL4-binding sites upstream of the adenovirus E4 TATA box¹⁰. Figure 2 shows the use of a gel retardation assay to measure the amount of three GAL4 derivatives required to fill their sites on each of the reporter templates used in a subsequent transcription experiment. The transcription experiment was then performed in HeLa cell nuclear extracts in the presence of excess activator; the product mRNAs were measured by primer extension analysis (Fig. 3a). Figure 3b summarizes the data shown in Fig. 3a. For all three activators, synergism was readily apparent (see, for example, the response of one versus nine sites using GAL4(1–147)+AH). Also as observed *in vivo*, as the number of GAL4 binding sites increases, the differences in response to a weak (GAL4(1–147)+AH) and the strong (GAL4(1–147)+VP16) activator greatly diminishes; at even greater numbers of sites the response to GAL4(1–147) approaches that seen with the other activators (data not shown).

A further demonstration that the experiment in Fig. 3a was done under conditions in which the activator binding sites were saturated is provided by the experiments of Fig. 3c and d. Figure 3c describes activation by increasing amounts of GAL4(1–147)+VP16 on templates bearing various numbers of binding sites and Fig. 3d shows the results of parallel DNase I footprinting

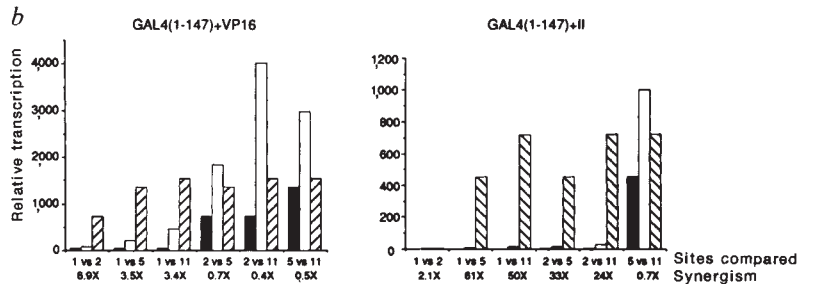
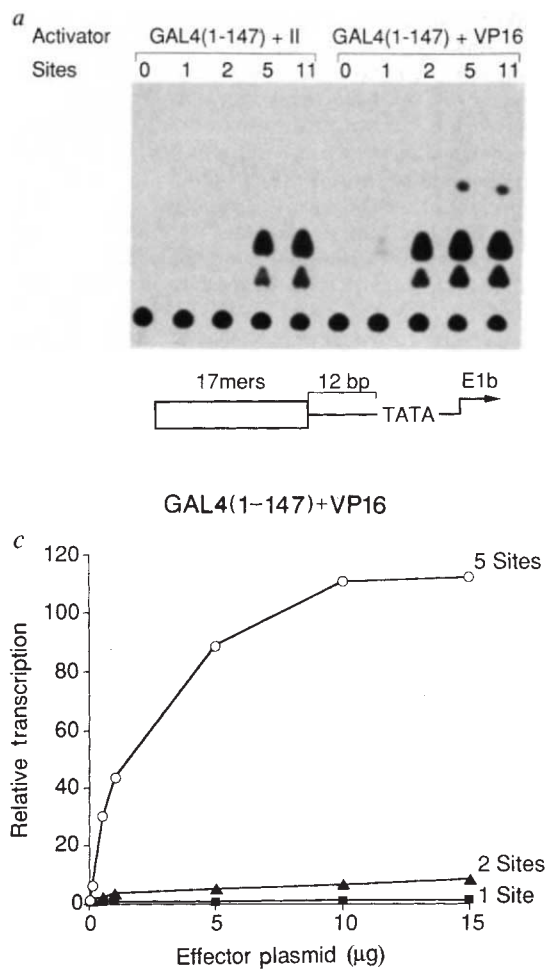
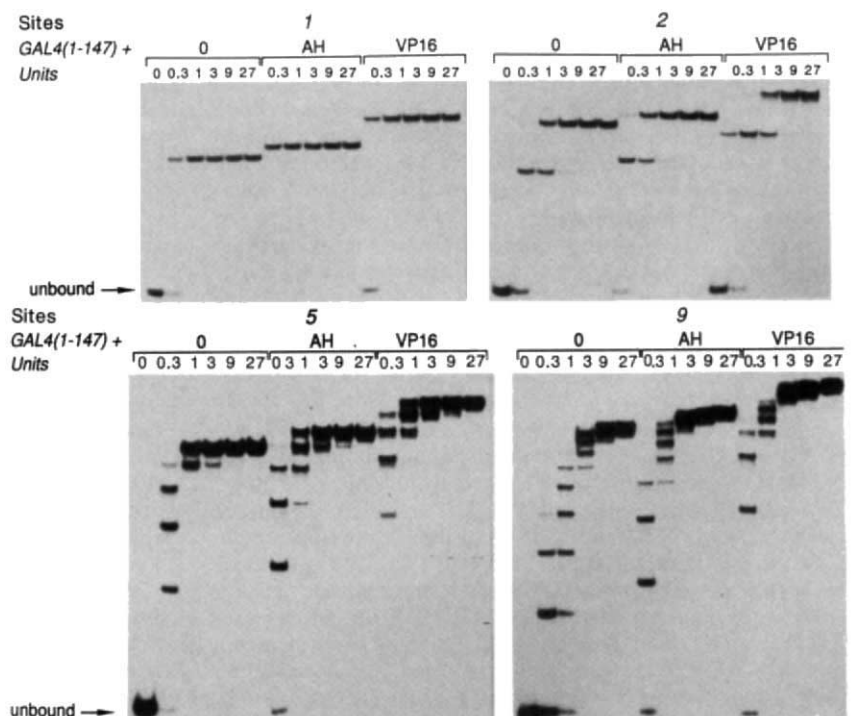


FIG. 1 Transcriptional synergism *in vivo*. *a*, Templates bearing the indicated number of 17-mers cloned 12 bp upstream of the E1b TATA box were transfected into CHO cells with effector plasmids directing the synthesis of GAL4(1-147) + II and GAL4(1-147) + VP16. After 60 h cells were collected and assayed for CAT activity. An autoradiogram of a typical experiment is shown. The template is indicated below. *b*, Bar graphs representing the data in *a*. Solid bars represent the amount of transcription found with the lower number of sites, hatched bars represent the amount found with the higher number of sites. The white bars are the expected value if the sites were having a simple additive effect. We compared each possible combination of sites, for example one versus two sites. The degree of synergism, indicated below each comparison, is defined as the ratio of the observed (cross hatched) to what is predicted if the sites were having an additive effect on transcription (white). *c*, A typical titration of effector plasmid. Increasing amounts of GAL4(1-147) + II effector plasmid were co-transfected into CHO cells with a constant amount of a reporter plasmid containing the indicated number of GAL4-binding sites. CAT activity was measured and plotted against the amount of effector DNA. The level of transcription in the absence of effector plasmid was assigned a value of 1.

METHODS. Transfections: 5 μ g of reporter plasmid and the indicated amounts of effector plasmid were transfected onto CHO cells using the DEAE-Dextran technique¹³ and CAT activity was measured and quantitated as described previously¹⁴. Cloning: the plasmid pE1bCAT (ref. 7) contains the adenovirus E1b TATA box cloned immediately upstream of the CAT gene. The reporter templates contained 1, 2, 5 and 11 tandem GAL4-binding sites cloned into the *Hinc*II site of pE1bCAT.

FIG. 2 Normalizing DNA binding of the activators to the reporter templates for *in vitro* transcription. The indicated GAL4 derivatives were titrated in reactions containing ³²P-labelled restriction fragments bearing the indicated number of GAL4-binding sites (17-mers). The conditions resembled those used in the *in vitro* transcription reactions but lacked HeLa nuclear extract. The resulting complexes were fractionated by electrophoresis through non-denaturing polyacrylamide gels as described previously⁸. The experiment in Fig. 3a was carried out under conditions in which the template sites were saturated with activator as determined by this assay. At low protein concentrations each activator generates the number of complexes expected from the number of template sites. The larger activators give rise to complexes with slower electrophoretic mobilities than the smaller activators.



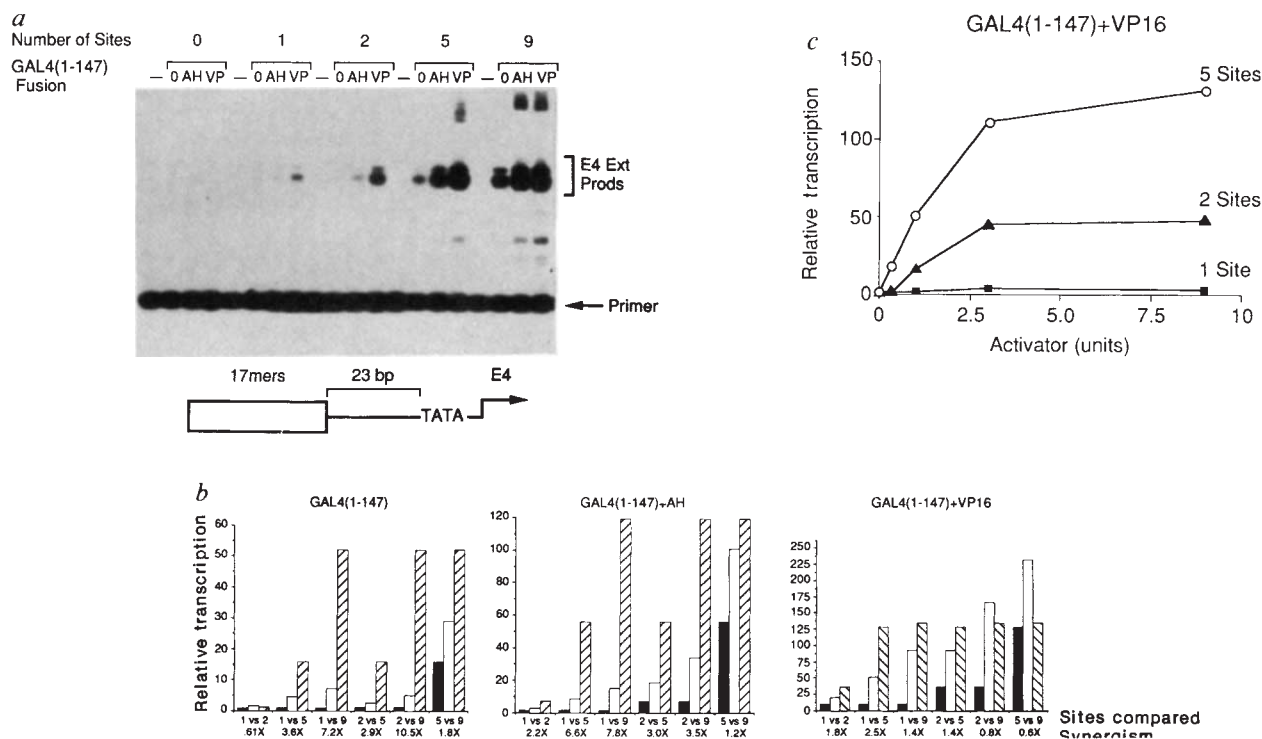
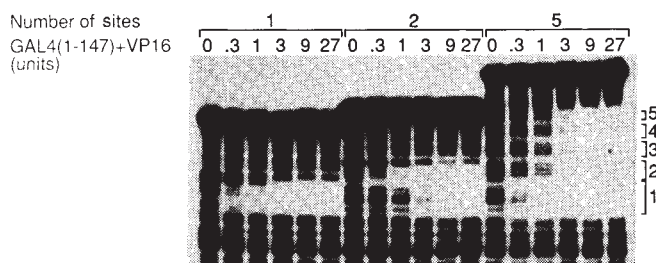


FIG. 3 Transcriptional synergism *in vitro*. **a**, Templates bearing the indicated number of 17-mers cloned 23 bp upstream of the E4 TATA box were transcribed in HeLa cell nuclear extract in the absence (—) or presence of 9 units of GAL4(1-147), designated 0, GAL4(1-147)+AH, designated AH, or GAL4(1-147)+VP16, designated VP. The reaction products were measured by primer extension¹⁰ and fractionated on a polyacrylamide-urea gel. An autoradiograph of the gel is shown. The primer and E4 extension products are indicated to the right of the autoradiograph. The multiple products are the result of transcription initiating at different points. **b**, The autoradiograph in **a** was scanned with a Helena densitometer, using dilutions of the primer to create a standard curve, and plotted as a bar graph. Each activator is represented by a separate graph. The background transcription in the presence of no sites was assigned a value of 1 and subtracted from the stimulated signal, the bars are described in the legend to Fig. 1. **c**, GAL4(1-147)+VP16 was titrated in transcription reactions on templates bearing the indicated number of binding sites. The autoradiographs were scanned with a Helena densitometer and plotted against the level of activator. The side-by-side comparison in **a** was made with 9 units of activator, an amount on the plateau level of the curve. **d**, DNase I footprinting



assays were performed in the nuclear extract in parallel to the transcription assays shown in **c**.

METHODS. Activator purification, *in vitro* transcription and DNase I footprinting assays have been described^{10,15,16}. Cloning: the plasmid pΔ-38 (ref. 10) contains the E4 TATA box, start site and 250 bp of coding sequence cloned into the *Sma*I site of pGem3. The plasmids pG₁E4T, pG₂E4T, pG₅E4T and pG₉E4T were constructed by inserting one, two, five and nine tandem GAL4-binding sites into the *Hinc*II site of pΔ-38, 23 bp upstream from the E4 TATA box (E4T).

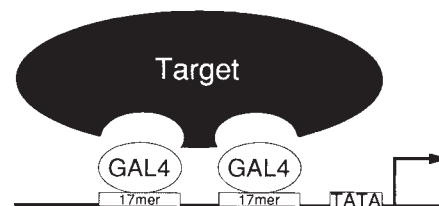


FIG. 4 Schematic representation of multiple activators simultaneously contacting a common target. Two molecules of a GAL4 derivative, bound to adjacent sites on DNA, simultaneously contact an as-yet-undefined target protein. The interaction energies would add, leading to an exponential increase in the affinity of the target for multiple activators; the exponential increase in affinity would generate a synergistic increase in transcription. Our experiments suggest that more than two, and possibly as many as 5–10 molecules of certain activators may simultaneously interact with the target. The target is shown as a single molecule although there may be multiple targets that interact with one another.

assays. The amount of protein added in the experiment of Fig. 3a is at least threefold above that required for maximal transcription and is sufficient to fill the template sites in a DNase I footprinting experiment.

We conclude that the GAL4-derived activators studied here can work cooperatively under conditions in which their sites on DNA are saturated. The accompanying paper⁴ shows that this statement can be extended to describe the synergistic effect of GAL4 and a mammalian transcriptional activator, ATF. These results are consistent with the idea that DNA-bound GAL4 molecules simultaneously contact a component(s) of the transcriptional apparatus as shown schematically in Fig. 4. This idea would explain the evidently promiscuous properties of various eukaryotic activators (see, for example, refs 11 and 12): any set of activator molecules that could simultaneously interact with some part of the transcriptional machinery would work cooperatively. If this idea is correct, our data suggest that somewhere between 5–10 molecules of activator can contact

the transcriptional target simultaneously. We note that according to this scenario the activator(s) might, like λ repressor, bind DNA cooperatively when present at sufficiently low concentration, but the effect would be mediated not by direct touching between the activators, but rather by indirect interaction. $\square$

Received 17 November 1989; accepted 23 March 1990.

1. Johnson, A. D., Pabo, C. O. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5061-5065 (1979).
2. Hochschild, A., Irwin, N. & Ptashne, M. *Cell* **32**, 319-325 (1983).
3. Giniger, E. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 383-386 (1987).
4. Lin, Y.-S., Carey, M. F., Ptashne, M. & Green, M. R. *Nature* **345**, 359-361 (1990).
5. Ma, J. & Ptashne, M. *Cell* **48**, 847-853 (1989).
6. Sadowski, I., Ma, J., Triezenberg, S. & Ptashne, M. *Nature* **335**, 563-564 (1988).
7. Lillie, J. W. & Green, M. R. *Nature* **338**, 39-44 (1989).
8. Carey, M., Kakidani, H., Leatherwood, J., Mostashari, F. & Ptashne, M. *J. molec. Biol.* **209**, 423-432 (1989).
9. Giniger, E. & Ptashne, M. *Nature* **330**, 670-672 (1987).
10. Lin, Y. S., Carey, M., Ptashne, M. & Green, M. R. *Cell* **54**, 659-664 (1988).
11. Davidson, I., Xiao, J. H., Rosales, R., Staub, A. & Chambon, P. *Cell* **54**, 931-942 (1988).
12. Strahle, U., Schmid, W. & Schutz, G. *EMBO J.* **11**, 3389-3395 (1988).
13. Lopata, M. A., Cleveland, D. W. & Sollner-Webb, B. *Nucleic Acids Res.* **12**, 5705-5717 (1984).
14. Lin, Y.-S. & Green, M. R. *Nature* **340**, 656-659 (1989).
15. Chasen, D., Leatherwood, J., Carey, M., Ptashne, M. & Kornberg, R. *Molec. cell. Biol.* **9**, 4746-4749 (1989).
16. Carey, M., Leatherwood, J. & Ptashne, M. *Science* **247**, 710-712 (1990).

ACKNOWLEDGEMENTS. This work was supported by grants from the national Institutes of Health and American Cancer Society to M.P. and M.R.G. M.C. was supported by a postdoctoral fellowship from the American Cancer Society.

CORRECTION

Muscle-specific gene expression controlled by a regulatory element lacking a MyoD1-binding site

Timothy J. Baldwin & Steven J. Burden

Nature **341**, 716-720 (1989).

WE reported that muscle-specific gene expression could be conferred by a *cis*-acting regulatory region in the murine acetylcholine receptor δ -subunit gene that lacks a MyoD1-binding site. Subsequently, we found that the sequence of several plasmid constructs used in these experiments was different from that reported. We reconstructed these plasmids, verified their sequence and, as studied previously, analysed their ability to confer muscle-specific expression on the *c-fos* basal promoter (FBP), a tissue-nonspecific promoter. We demonstrated previously that nucleotides -148/-108 of the δ -subunit gene are necessary for muscle-specific expression and bind myotube nuclear proteins distinct from MyoD1. We concluded that the MyoD1-binding site (-23/-15) is not sufficient to confer muscle-specific expression, and that an activity distinct from MyoD1 could bind to a *cis*-acting region necessary for muscle-specific expression. We confirm these results. We also reported that the FBP alone conferred low-level expression in C2 myotubes and 3T3 cells, whereas -148/-95/FBP and -148/-53/FBP conferred high level expression only in C2 myotubes. Therefore we concluded that -148/-95, a *cis*-acting region that lacks a MyoD1-binding site, was sufficient to confer muscle-specific expression upon a tissue-nonspecific promoter. We now find that the FBP alone and -148/-95/FBP confer identical low-level expression in C2 myotubes and 3T3 cells. In addition, -148/-53/FBP confers high-level expression not only in C2 myotube cultures, but also in 3T3 cells. Consequently, we have no evidence that the *cis*-acting regulatory region that lacks a MyoD1-binding site (-148/-95) is sufficient to confer muscle-specific expression. $\square$

nature
is available in
microform.



University Microfilms

International reproduces this publication in microform: microfiche and 16mm or 35mm film. For information about this publication or any of the more than 13,000 titles we offer, complete and mail the coupon to: University Microfilms International, 300 N. Zeeb Road, Ann Arbor, MI 48106. Call us toll-free for an immediate response: 800-521-3044. Or call collect in Michigan, Alaska and Hawaii: 313-761-4700.

☐ Please send information about these titles:

Name _____

Company/Institution _____

Address _____

City _____

State _____ Zip _____

Phone () _____

University
Microfilms
International

Antibody-directed drug discovery

M.E. Wolff and A. McPherson

By harnessing the remarkable power of the immune system to access the three-dimensional chemical information sequestered in target sites for drug ligands, antibody-directed drug discovery speeds the design of new therapeutic entities.

FOR more than a century, the search¹ for new drugs has been rooted in an organic chemistry-driven technology based on the synthesis and biological evaluation of large numbers of candidate compounds. Highlighting the inefficiency of this process is the estimate in *Bio/Technology* that US pharmaceutical companies spent about \$14,000 million in 1989 to achieve approvals of only 22 new chemicals. On average, 10,000 candidates were prepared and tested for each successful drug.

The growing costliness and lengthiness of traditional methodology has focused renewed attention on what has been called rational drug design — the ability to create pharmaceuticals on the basis of a detailed knowledge of their structure and their interaction with biological targets. Today, most major pharmaceutical companies and several start-up companies have organized efforts in rational drug design. Several distinct strategies are used, including approaches based on computer modelling² of data obtained from binding assays and enzyme crystallography, as well as methods depending on bio-organic chemistry³. These applications of rational drug design thus require at least the purification and sometimes the crystallization of targets that are central to each disease process. Yet many prospective drug targets are membrane-bound macromolecules that cannot readily be crystallized or even identified.

Typical of the techniques being developed to tackle these difficulties are those of ImmunoPharmaceutics, under the term 'antibody-directed drug design' (Fig. 1). The remarkable power of the mammalian immune system is first harnessed to access the three-dimensional structural information contained in pharmacological receptors, enzymes and viruses without the need for special purification or isolation efforts. Then, the pharmacophoric information embedded in the resulting monoclonal antibody is read out by one of several methods and finally used to construct a peptidomimetic drug. These steps are considered briefly in the next sections.

Antibodies to drug targets

Through monoclonal technology, idiotypic antibodies (Ab-1) to key drug action sites in receptors, enzymes and viruses can be selected and produced in any desired amount without the need to isolate the target (Fig. 2). The idiotypic antibody is a

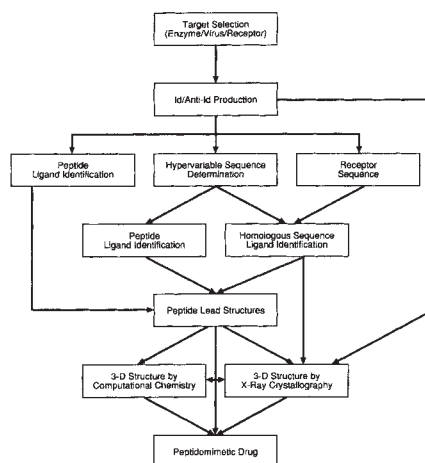


FIG. 1 Technologies for antibody-directed drug design.

mirror image of the target site, both in a spatial and electronic sense. By using the idiotype antibody as an antigen, an anti-idiotypic antibody (Ab-2)⁴, can be produced as a positive image of the target site (see Fig. 2).

Important drug design information is obtained from such antibodies or their fragments or complexes by crystallization using conventional protein crystallization techniques⁵. After collection of X-ray diffraction data the crystal structure is solved using the conventional multiple isomorphous replacement technique requiring heavy atom derivitization of the crystals. An alternative approach is that of molecular replacement, which depends on some degree of knowledge of the structure of the antibody, the antigen, or both.

As an example, the three-dimensional structure⁶ of an antilysozyme Fab-lysozyme complex is shown in Fig. 3. A close-up view (Fig. 4) shows the spatial and electronic complementarity of the interacting surfaces of lysozyme residues 41–46 and

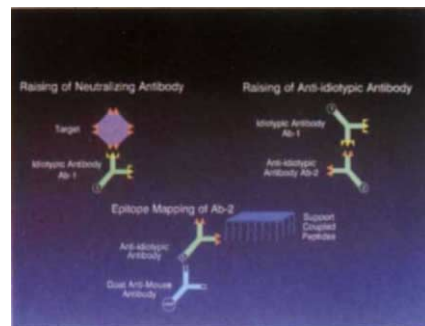


FIG. 2 Initial steps in antibody-directed drug design.

light-chain residues 55–59 that contributes to the hydrophobic, electrostatic, and hydrogen bonding accounting for the nanomolar affinity of the two ligands.

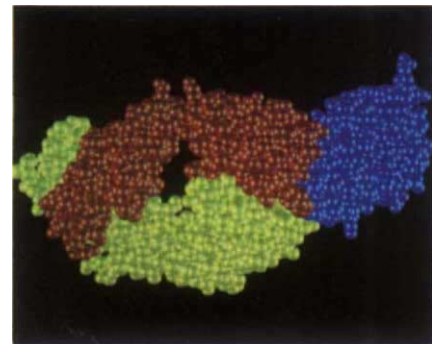


FIG. 3 Monocolor-coded CPK model representation of lysozyme-HyHEL-5 complex⁶ obtained by Drs V.N. Balaji and U.C. Singh on an Evans and Sutherland PS 390 graphics system using MOGLI graphics display software and coordinates in the Protein Data Bank. Blue atoms: lysozyme; red atoms: light chain; and yellow atoms: heavy chain.



FIG. 4 Monocolor-coded stick model close-up view of the contact area of Fig. 3. Contact residues (<4 Å) are highlighted and shown as solvent-accessible surface.

These specific interactions between antibody and antigen demonstrate dramatically the potential of the approach for directed drug design. Such analyses provide an unequalled structural basis for the creation of synthetic molecules, peptides, or peptidomimetics that share with the idiotype antibody the ability to interact with the antigen. Related information can be deduced for anti-idiotypic antibodies that are replicas of receptors or other inaccessible macromolecules.

Extraction of information

In addition to X-ray crystallography, three other techniques (see Fig. 1) are used to read out structural information

from the embedded pharmacophore in the antibody or from the anti-idiotypic antibody. By using the polymerase chain reaction, the amino acid sequence of the hypervariable loops of the antibodies can be determined. Alternatively, peptides that are complementary to the surface of anti-idiotypic receptor surrogates can be synthesized using solid-phase methods (see Fig. 2)⁷. A properly selected peptide that is complementary to a receptor surrogate results in a receptor antagonist. Lastly, evidence for the location of key amino acid sequences in the hypervariable loops of anti-idiotypic antibodies can be added from their homology with sequences in the initial antigen⁸.

Peptidomimetic synthesis

The preceding process results in the design of a peptide drug that may be an enzyme inhibitor, a receptor antagonist (or sometimes an agonist), or an antiviral. Because peptides typically have short biological half-lives and poor oral bioavailability, it remains to synthesize a peptidomimetic⁹ based on the peptide structure that is biologically active and can be administered orally.

The design of such peptidomimetics requires two different applications of computer-assisted drug design (CADD) techniques (V. N. Balaji and U. C. Singh, personal communication). First, if an estimate of the bioactive conformation of the peptide is not available from X-ray or nuclear magnetic resonance, data must be obtained from primary sequence information through the application of computational chemistry algorithms^{10,11}. Based on the predicted three-dimensional spatial and electronic characteristics of the peptide, a proprietary database and algorithm of peptidomimetic surrogate elements is used to design candidate drugs for synthesis. Docking routines may be used to optimize complementarity in those cases where the three-dimensional structure of the receptor surrogate has been obtained^{10,11}.

Antibody-directed drug design offers wholly new insights and approaches in the design and optimization of new compounds for therapeutic application, and risk is minimized by the multiplicity of available strategies. As these methods are developed further, even greater savings in time and effort will be at hand in the search for new drugs. □

Manfred E. Wolff is Vice President for Research and Development, ImmunoPharmaceuticals, Inc., 11011 Via Frontera, San Diego, California 92127, USA, and Alexander McPherson is Professor and Chairman of the Department of Biochemistry, University of California, Riverside, USA. For more information, fill in reader service number 100.

1. Burger, A. in *Burger's Medicinal Chemistry* 4th Edn, Pt 1. (Ed. Wolff, M.E.) 1–54 (Wiley, New York, 1980).

Immunology roundup

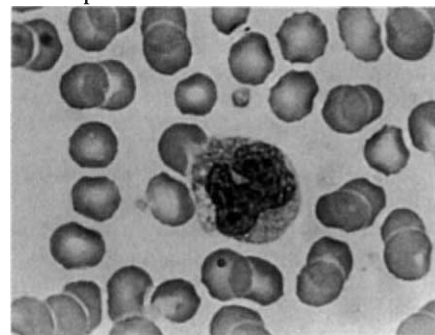
A colour processing system for gold-tagged antibodies, a tumour necrosis factor α [¹²⁵I] RIA kit, and a whole host of monoclonal antibodies are among this week's new products for immunological research.

As a follow up to its current line of human MAPPING (message amplification phenotyping) Amplimers, Clontech has issued new mouse cytokine MAPPING Amplimers for **cytokine analysis** using the GeneAmp PCR process (*Reader Service No. 101*). Designed to be faster and more sensitive than conventional methods such as immunoassays, and without using radioactivity, the MAPPING process involves the isolation and reverse transcription of mRNA and the specific PCR amplification of cytokine cDNAs. Each MAPPING amplimer set contains a pair of primers for a specific cytokine and a detailed protocol. Amplimers are sold under licence from Perkin-Elmer Cetus and produced in collaboration with Genelabs.

Eastman Kodak has developed a complete **colour processing system** — the Gold Plus enhanced immunogold staining kit — for the visualization by light microscopy of gold-tagged biological specimens (*Reader Service No. 102*). First, specific antigens are localized with gold-tagged antibodies. Then, silver halide is reduced in the presence of hydroquinone and adsorbed onto the colloidal gold surface, amplifying the signal. Following the silver-enhancement step, the specimens are placed in a stop bath and fixed. The biological film emulsion is then dip coated using the Gold Plus reagents: cyan, magenta or yellow. The colour processing system uses the silver-enhanced gold tag as the catalytic centre for the colour coupler/dye reaction, with colour development taking 3 to 60 seconds, says Kodak. Processing options are available further to enhance the colour intensity in accordance with the signal desired and the sample background. Colour signals are particularly useful when working with thick tissue sections and with structures having high inherent density. Each Gold Plus kit contains sufficient reagents for 50 preparations.

For the **one-step separation of monocytes** from leucocyte-rich plasma and whole blood, Nycomed recommends its

specialized formulated Nycodenz monocytes solution (*Reader Service No. 103*). Traditional methods of monocyte purification rely on the adherence properties of cells and, consequently, a significant proportion of cells are discarded, says Nycomed. Nycodenz can be used to isolate all classes of monocytes. The solution provided is slightly hypertonic: this causes shrinkage of lymphocytes, which pellet to the bottom of the tube



For the one-step isolation of monocytes, use Nycodenz from Nycomed.

while the monocytes remain at the top of the gradient. Full separation is achieved by overlaying Nycodenz with whole blood and centrifuging at 600 g for 15 minutes. Nycodenz is supplied sterile and in ready-to-use 100-ml quantities. When kept sterile and protected from exposure to light, the solution is stable for up to five years.

Promega has a new complement of **extracellular matrix products** that should be of particular interest to cell culture researchers (*Reader Service No. 104*). E-C-L cell attachment matrix, which is rich in enactin, laminin and collagen type IV, promotes the attachment and spreading of many anchorage-dependent cell types, says Promega. It is designed either to coat culture substrates or be incorporated into culture media. On the other hand, human fibronectin, Promega's plasma-derived protein, can be used in cell culture to augment the attachment of a variety of cell types under reduced or serum-free conditions. Human fibronectin has high purity and is virally inactivated, says Promega.

- Desjarlais, R.L., Sheridan, R.P., Seibel, G.L., Dixon, J.S. & Kuntz, I.D. *J. med. Chem.* **31**, 722–729 (1988).
- Walsh, C.T. *Trends biochem. Sci.* **8**, 254–257 (1983).
- Anti-idiotypes, Receptors, and Molecular Mimicry (Eds Linthicum, D.S. & Farid, N.R.) 1–154, 255–284 (Springer, New York, 1988).
- McPherson, A. in *The Preparation and Analysis of Protein Crystals* (Wiley, New York, 1982).
- Sheriff, S. et al *Proc. natn. Acad. Sci. U.S.A.* **84**, 8075–

- 8079 (1987).
- Geysen, H.M., Rodda, S.J., Mason, T.J., Tribbick, G. & Schoofs, P.G. *J. immunol. Meth.* **102**, 259–274 (1987).
- Williams, W.V. et al *Proc. natn. Acad. Sci. U.S.A.* **85**, 6488–6492 (1988).
- Goodman, M. *Biopolymers* **24**, 137–155 (1987).
- Kuntz, I.D. *Protein Eng. (England)* **1**, 147–148 (1987).
- Cohen, F.E., Gregoret, L., Presnell, S.R. & Kuntz, I.D. *Prog. clin. Biol. Res.* **289**, 75–85 (1989).

Lastly, Promega's rabbit, anti-human fibronectin can be used to localize fibronectin by immunological means.

Metachem Diagnostics is the UK distributor of Advanced Magnetics' **tumour necrosis factor α (TNF α) [125 I] RIA kit**, which can be used to measure levels of natural and recombinant TNF α in serum, citrated or heparinized plasma samples, or cell culture supernatants (*Reader Service No. 105*). The solution-phase RIA is based on the 'sandwich' principle and uses an 125 I-labelled monoclonal antibody (mAb) and a mAb that binds to TNF α as well as to the solid phase. Incubation of TNF α with the two mAbs results in the formation of a complex where TNF α is sandwiched between the antibodies. The complex is separated from solution by magnetic separation or centrifugation with magnetic second antibody, which binds the non-isotopic monoclonal portion of the complex. The magnetic pellet that forms is counted in a gamma counter and the TNF α concentration determined by interpolation from a standard curve. The RIA has a sensitivity of 3.08 pg 0.1 ml^{-1} , is highly specific, and has negligible cross reactivity, says Metachem. Each kit contains sufficient reagents for 100 sample tubes including standard curves.

Cytokines: Il-8

The latest cytokine from the Bachem stable is **endothelial interleukin-8 (Il-8)**, a 77 amino acid variant of monocyte Il-8 with a pentapeptide extension at the amino terminus (*Reader Service No. 106*). According to Bachem, endothelial Il-8 has been shown to promote neutrophil chemotaxis and degranulation, and to inhibit neutrophil binding to endothelial cells in the presence of inflammatory stimuli. As the adhesion inhibitory activity helps to limit neutrophil-mediated endothelial damage during inflammation, Il-8 may provide new insights into natural anti-inflammatory mechanisms. Endothelial Il-8 is produced as a fusion protein in *Escherichia coli* and is >98 per cent pure by HPLC, says Bachem. It is fully active in chemotactic assays involving lymphocytes and neutrophils.

Boston-based Endogen has introduced a polyvalent **antibody to human interleukin-8 (Il-8)** — the novel cytokine that has been shown to cause activation and chemotaxis of neutrophils *in vivo* (*Reader Service No. 107*). Anti-Il-8 is purified from the serum of rabbits immunized with the native sequence recombinant human cytokine. It is supplied sterile and free of carrier proteins and preservatives, says Endogen. Applications for the antibody include: use as a specific antibody reagent in ELISAs and radioimmunoassays; neutralization experiments; immunocyto-

chemical staining; radioimmunoprecipitation; and immunoprecipitation in Western blots, immunoblots and slotblots. Endogen can also supply the Il-8 antibody coupled to a gel matrix. In this form, the pre-coupled antibody is supplied complete with buffers and a reusable column, and is suitable for use in cytokine purification, immunoprecipitation and neutralization procedures.

The cytokine **interleukin-8 (Il-8)** and the antibody anti-Il-8 are two new products from Janssen Biochimica (*Reader Service No. 108*). Il-8, which is produced by monocytes, fibroblasts, T cells, endo-



Il-8 and rabbit, anti-Il-8: now available from one source — Janssen.

thelial cells, dendritic cells and keratinocytes, is known to stimulate neutrophil and T-cell chemotaxis, activate leucocytes and augment neutrophil lysosomal enzyme release, superoxide generation and antibacterial and antifungal activities. In excessive concentrations, Il-8 can cause plasma leakage and can increase vascular permeability. In addition, high levels of Il-8 are associated with pathological states — psoriasis, rheumatoid arthritis and asthma. Completing the picture, is Janssen's polyclonal rabbit, anti-human Il-8, which is specific to Il-8, does not cross react, and recognizes both natural and recombinant forms, says Janssen. Each vial of anti-Il-8 contains 1 mg ml^{-1} , and is free of preservatives and carrier proteins.

Cytokines: Il-2

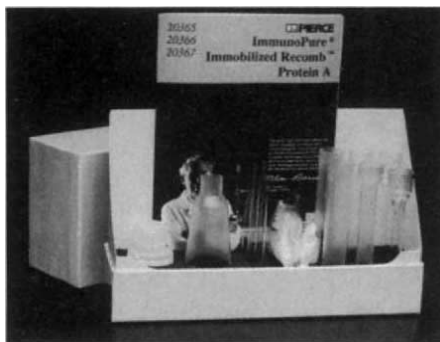
Pharmacia Diagnostics' **natural, human interleukin-2 (Il-2)** is new to the marketplace and is designed to provide an economical alternative to producing, purifying and testing Il-2 in the laboratory (*Reader Service No. 109*). Natural Il-2 stimulates the growth of a variety of T-cell types and clones and is suitable for both human and murine studies, says Pharmacia. Each lot of natural Il-2 has a minimum potency of 640 units per ml and is stable for an indefinite period when stored at -20°C . Pharmacia's Il-2 is supplied sterile in ready-to-use 10- and 50-ml quantities complete with data sheet.

Dynal has developed a rapid **bioassay for Il-2** that does not require the maintenance of Il-2-dependent cell lines (*Reader Service No. 110*). The assay uses Dynabead M-450 magnetizable polymer particles

coated with anti-CD3 to isolate T cells that are activated to express Il-2. CD3-positive T cells isolated in this way are, says Dynal, free from accessory cells and, as such, express Il-2 receptors but do not produce Il-2. The presence of exogenous Il-2 causes these cells to exhibit a proliferative response, which is Il-2 dose-dependent and can be measured by the incorporation of tritium-labelled thymidine. The assay has a sensitivity of $0.01 \text{ U Il-2 ml}^{-1}$, which is comparable to existing methods that rely on cells obtained from long-term culture. It has a dose-response curve that is linear up to at least $20 \text{ U Il-2 ml}^{-1}$, says Dynal.

Protein binding

Pierce now offers **immobilized recombinant Protein A** with a leak-resistant linkage (*Reader Service No. 111*). This recombinant form of Protein A immobilized on a cross-linked agarose matrix prevents the mitogenic effects that can occur with native Protein A contamination of samples, says Pierce. The recombinant Protein A is harvested from a non-pathogenic form of *Bacillus* that has been

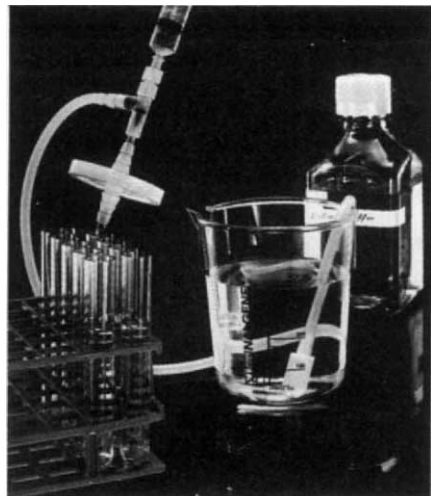


Pierce's immobilized Protein A with leak-resistant linkage.

genetically engineered to secrete the protein. Consequently, recombinant Protein A is enterotoxin-free. Immobilized recombinant Protein A is compatible with Pierce's ImmunoPure (A) IgG buffer system and its gentle Ag/Ab buffer system.

Genex has eight new **GammaBind Protein G conjugates** designed to provide increased sensitivity and lower background interference with Western blots or ELISAs (*Reader Service No. 112*). GammaBind G is a genetically engineered Protein G that binds only IgG from a broad range of mammals, including all subclasses of human, mouse, rat, goat, sheep, cow, horse and rabbit. Genex has made available GammaBind G-gold for blotting purposes and GammaBind G-HRP, FITC and alkaline phosphatase for ELISAs. Completing the line up is GammaBind G biotin with three streptavidin conjugates for ELISAs and immunohistological studies.

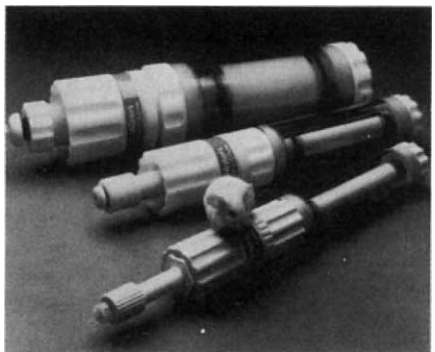
Pre-bound recombinant **Protein G antibody purification devices** are now available from Nalge (*Reader Service No. 113*). Protein G has been shown to be useful for immunoglobulin purification, especially with murine monoclonal antibodies. It has a higher selectivity than Protein A for human, mouse or other species'



Pre-bound rProtein G antibody purification devices from Nalge.

classes of IgG and shows no cross reactivity with IgM, IgE, IgA or IgD, says Nalge. Scale up is made possible by stacking multiple 25- or 47-mm matrix discs in a holder, or by using Nalge's pre-stacked 50-mm units. Typically, recombinant Protein G-bound 50-mm units have a maximum yield of 16 mg of human IgG.

Protein-Pak HR high-performance **ion exchange glass columns** from Waters are designed for the analytical and preparative separation of immunoglobulins from ascites or cell culture supernatants (*Reader Service No. 114*). Packed with 8-, 15- or 40- μ m particle size anion (DEAE) or cation (SP) exchange resins (1,000 Å), the Protein-Pak HR columns provide improved resolution, protein binding capacity and recovery, says Waters. Scale up is simplified by the use of different column dimensions: 10, 20, or 50 mm $\times$ 100 mm. When used in combination with Waters' 650 advanced protein purification



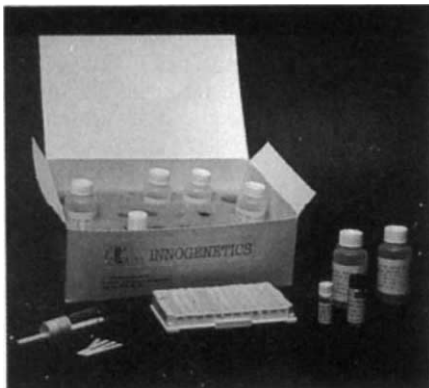
Waters' Protein-Pak ion exchange glass columns come pre-packed.

system, performance of the Protein-Pak columns is optimized.

Isotyping kits

The new monoclonal antibody (mAb) panel from Silenus Laboratories of Australia provides researchers with a direct method of **serotyping *Neisseria gonorrhoeae*** (*Reader Service No. 115*). The ten mAbs react specifically with the epitopes of protein I of *N. gonorrhoeae* and can be used for the direct detection of *N. gonorrhoeae* in patient specimens, or for colony confirmation or research applications. Silenus sells the mAbs individually or as a cocktail of ten mAbs that will react with 98 per cent of serovars worldwide. In both forms, they can be supplied either unconjugated or conjugated with FITC.

With the INNO-LIA mouse monoclonal antibody (mAb) **isotyping kit**, all murine subclasses — IgG1, IgG2a, IgG2b, IgG3, IgM, IgA and IgE — can be detected in less than one hour, says Belgium-based Innogenetics (*Reader Service No. 116*). The protocol differs from conventional ELISA techniques in that rat mAbs against the different subclasses are applied in fine lines to a membrane-coated strip. One test strip is used per sample.



Innogenetics' INNO-LIA mouse isotyping kit for speed and sensitivity.

Using monoclonals for the purposes of detection eliminates problems of cross reactivity, says Innogenetics, and provides a sensitivity of 0.1 μ g ml⁻¹. Results are read visually and the strips can be stored as a permanent record.

Oncoproteins

Adding to its list of over 2,000 immunological reagents, Bioproducts for Science has introduced a new line of **oncoproteins and antibodies to oncoproteins** (*Reader Service No. 117*). Oncoprotein antibodies are site-specific probes raised against small synthetic peptides that are selected from a particular oncoprotein's amino acid sequence. They can be used to recognize identical sequences in intact native proteins as well as the synthetic peptide. The antibodies have been characterized against both the synthetic peptide and the native protein released from cells express-

ing the relevant oncoprotein, says BPS, and are supplied as purified immunoglobulin complete with test data and recommended working dilutions. Stated applications for the antibodies include immunoprecipitation, immunoblotting, immunohistochemistry, affinity purification and neutralization studies.

Capitalizing on Cambridge Research Biochemicals' expertise in the synthetic peptide approach to antibody production, the company offers a range of **murine hybridomas secreting antibodies** capable of specifically targeting epitopes contained within the cytoplasmic and external domains of both human epidermal growth factor-R (EGF-R) and *c-erbB-2*, and from the nucleus-associated *myc* oncoprotein family (*Reader Service No. 118*). All have been characterized for native protein reactivity by immunoblotting, flow cytometry, immunoprecipitation and immunocytochemistry. In view of the importance of reagents specific for EGF-R, *c-erbB-2* and *myc* proteins in the diagnostic and prognostic assessment of various abnormal physiological states, CRB's antibodies may have considerable value in clinical, diagnostic and therapeutic research. The hybridomas are available for licensing or purchase.

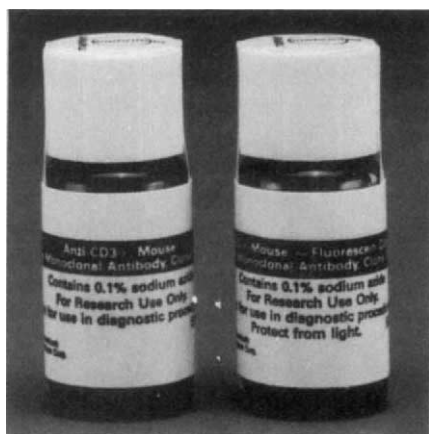
Monoclonal assortment

The latest addition to BioGenex's panel of antibodies for the identification of lymphoma and leucocyte markers on paraffin-embedded tissues includes a **monoclonal antibody to human macrophage** — LN-5 (*Reader Service No. 119*). LN-5 stains the cytoplasm of macrophages and histiocytes in haematopoietic organs and their derived tumours, says BioGenex. It is strongly reactive with cases of true histiocytic lymphoma, but negative, except in macrophages, in Hodgkins and non-Hodgkins lymphomas. LN-5 is available as a ready-to-use reagent, or as a complete kit with the StrAviGen super-sensitive detection system. The kit provides sufficient reagents for up to 50 slides.

Identi-T CyM1 is T Cell Sciences' latest **monoclonal antibody to the human T-cell antigen receptor (TCR)** (*Reader Service No. 120*). The mAb identifies a framework (or common) determinant of the γ chain of the human $\gamma\delta$ TCR. The determinant can be detected on fixed cells and frozen tissue sections, but is not intended for flow cytometry analysis as the binding determinant is not exposed on the surface of live cells, says T Cell Sciences. Identi-T CyM1 has applications in immunohistology protocols, immunoprecipitation procedures and Western blot determinations to investigate the role of T lymphocytes in lymphoid malignancies, autoimmune diseases and immune system disorders. The anti-

body is available in a ready-to-use purified form — sufficient for 100 tests.

The monoclonal antibody **anti-mouse CD3- ϵ** , an important tool in the immunologist's armoury for cellular studies involving the mouse model, is now available from Boehringer Mannheim (*Reader Service No. 121*). The anti-mouse CD3- ϵ , is a marker for targeting mature mouse T



Boehringer's anti-mouse CD3- ϵ simplifies T-cell mouse studies.

cells expressing the T-cell receptor. It can also be used directly to activate T cells, independent of antigen binding. The antibody is supplied in a purified form and is function tested by flow cytometry. For simple, one-step labelling techniques, anti-mouse CD3- ϵ is available conjugated to fluorescein.

New in the immunological marketplace is Serotec's **mouse anti-rat natural killer (NK) cell monoclonal antibody**, which recognizes and reacts with all rat NK cells, large granular lymphocytes and neutrophils (*Reader Service No. 122*). The anti-

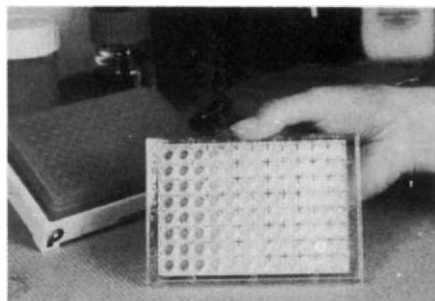


Serotec's mouse anti-rat natural killer cell monoclonal antibody.

body is available in three forms: 5-ml vials of culture supernatant containing $5 \mu\text{g ml}^{-1}$; 0.25-ml vials of ascites containing 1–2 mg ml^{-1} ; and as a F(ab')₂ FITC-conjugate in 1-ml vials containing $700 \mu\text{g ml}^{-1}$ of antibody.

Everything for ELISAs/RIAs

Polyfiltronics has a new line of **microtitre plates** called Filtraplates, which have 96 open-bottomed wells, each of which is sealed with a membrane that acts as a

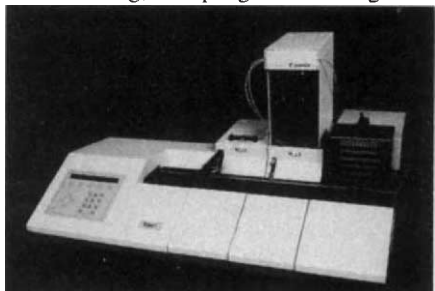


Polyfiltronics' Filtraplates for immunoassays and filtration purposes.

solid-phase support (*Reader Service No. 123*). Filtraplates can be equipped with membranes of various pore sizes and characteristics: high protein binding membranes for immunoassays, and low protein binding membranes for filtration purposes. According to Polyfiltronics, the patented encapsulation system of the plate eliminates cross-talk between wells. Liquid flowthrough can be augmented by using the Filtraplate in conjunction with Polyfiltronics' vacuum manifold system. The company also supplies activated 96-well microplates that have been chemically treated for the covalent attachment of proteins.

Wellfill 4 from Denley Instruments functions as both a single- and multi-reagent **dispenser for microtitre plates** (*Reader Service No. 124*). Suitable for enzyme-based or radioimmunoassays, Wellfill 4 can dispense up to eight different reagents in the range of 10 to 400 μl . Fixed volumes for routine applications are selected with plug-in cards. In addition, a fully programmable card is available for user control of dispense volumes that can be set to dispense in 1- μl increments. All parts of the Wellfill 4 dispenser that are in contact with liquid can be removed and autoclaved.

Dynatech now offers an **automatic ELISA processor** — the System 7000 — that is based on the company's MR 7000 automatic microplate reader (*Reader Service No. 125*). When functioning as a stand-alone automatic reader, the MR 7000 can read a plate in about 1.7 seconds, says Dynatech. Users can also choose from several optional features, including temperature control, UV operation, barcode reading, and program cartridges for

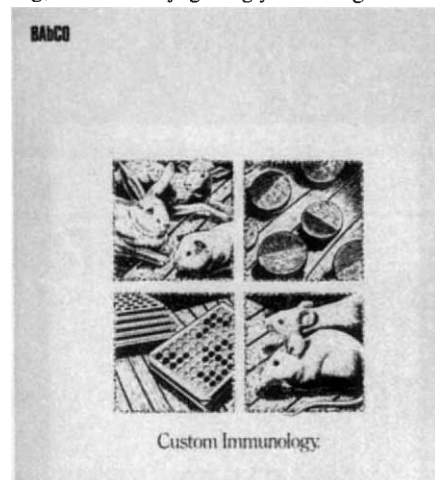


Dynatech's System 7000 automatic ELISA processor with modular design.

agglutination reading, kinetic analysis and tissue culture reading applications. By virtue of its modular design, the MR 7000 can be expanded to give full ELISA processing potential. Various add-on modules are available for microplate washing, dispensing and diluting, and random-access plate storage. The system's advanced plate transport system allows the single-plate carrier to travel from the reader to any number of the add-on modules up to a distance of two metres.

Newsbeat

Take the effort out of antibody and immunoassay development with Berkeley Antibody Company's (BAbCO's) **custom antibody and immunoassay development services** (*Reader Service No. 126*). BAbCO will generate polyclonal or monoclonal antibodies against peptides, proteins, organic haptens and other molecules to order. Also available are custom antigen preparation services for isolating, purifying, and/or conjugating your antigen. The



BAbCO's catalogue of custom antibody and immunoassay development.

line-up of other services includes: polyclonal antisera production in small and large animals, custom hybridoma development, scale-up of mAb production in ascites fluid, and immunobiochemical procedures that include antibody conjugations, purifications, fragmentation, biotinylation, isotyping, enzyme and fluorescent labelling. BAbCO also provides immunoassay development services using ELISA, PCFIA and Western blot formats.

Materials and services pertinent to cell physiology, virology, oncology and immunology can be found in the **1990 catalogue** from Lee and BioMolecular (*Reader Service No. 127*). The extensive listing includes interferons, animal viruses, antiviral sera and immunoglobulins. Many of the new products are particularly applicable to the development of vaccines and diagnostic products. Also outlined are research services for the identification and quantification of viruses, interferons, anti-interferon and antiviral

ADVERTISEMENTS

Custom DNA
Purified and Shipped in 48 hours,
Worldwide.

100 ug, \$10.00 a base for the first 15 bases,
\$7.50 for each additional base.
Delivered.

Research Genetics
1-800-533-4363
In the United Kingdom 0-800-89-1393
In Canada 1-800-533-4363
FAX 205-536-9016

Reader Service No.26

SPECIAL PRICE
CUSTOM DNA
100 μ g, 58.00 FF A BASE
FOR THE FIRST 15 BASES.
49.00 FF, each additional base
BIOPROBE SYSTEMS
CALL : 33-1-48-51-66-22
FAX : 33-1-48-51-59-90

Reader Service No.23

SINGER
INSTRUMENTS



**COMPLETE SYSTEMS
FOR
XENOPUS OOCYTE INJECTION**

INT. SALES: +44. 984. 40226.
Singer Instrument Co. Ltd.
Roadwater, Watchet, Somerset,
England, TA23 0QL. Tel: Washford (0984) 40226.
Telex: 337492 Corncab G

Reader Service No.50

WANTED
Ideas for Software Development
CHEMISTRY BIOCHEMISTRY
BIOTECHNOLOGY MEDICINE

I develop and distribute software for the above mentioned fields. I am searching for partner who can contribute product ideas and know-how. They have no further obligations and can participate on profit. If wished, a further going engagement is possible.
I have experience with Realtime-UNIX (ONX), HP-UX, DOS, UNIFY and C.
M. Weidmann — Technische EDV-Systeme
Oberhofstrasse 41, CH-6020 Emmenbrücke, Switzerland
Telephone: 041-559-329 Fax: 041-559-329

Reader Service No.21

immunoglobulins, sterilization of biologicals and *in vitro* toxicity testing.

More than 300 new immunological reagents and kits for use in research, clinical and diagnostic applications are featured in the latest **catalogue** from The Binding Site (*Reader Service No. 128*). The new product listings include IgG subclass reagents in kit form for the nephelometric assay of subclass levels, and a range



Product news from The Binding Site.

of purified myeloma proteins; five EIA kits for the investigation of functional IgG activity; and antisera and conjugates to mouse and rat proteins of a higher titre than previously offered.

Outlined in the FisherBiotech **ELISA systems brochure** are a series of microplate readers, providing varying levels of automation to suit laboratory needs (*Reader Service No. 129*). A range of equipment and supplies are featured in the brochure: from kinetic microplate readers, automated washers and software to immunological accessories such as microplates and reagents. Details of the latest ELISA data management system are included, which combines an IBM PS/2 PC with Fisher's Kineti-Calc software and provides full instrument control and data handling capabilities.

Pierce and Warriner says its 340-page full colour **immunotechnology catalogue and handbook** features over 200 new immunochemical products (*Reader Service No. 130*). This latest catalogue includes product information, protocols and references on the following subject areas: immunogens, antigen/antibody purification, antibodies, avidin-biotin, and protein modification.

Immunologists can now obtain the 1990 Jackson **immunoresearch catalogue** from Stratch Scientific (*Reader Service No. 131*). New product lines include affinity-purified antibodies and proteins, a range of phycoerythrin conjugates, Affinipure antibodies for double or multiple labelling and AMCA conjugates.

Immunoreagents/assays

Novabiochem recommends Liposorb for the **selective removal of lipoproteins** from blood plasma or serum, while allowing the antibody content of antiserum or the coagulation factors to remain in tact (*Reader Service No. 132*). Liposorb can be used as a dry substance, in batch procedures or by column. Novabiochem says Liposorb has low solubility in solvents, providing, says Novabiochem, stability in acids and bases. Liposorb is thermally stable at high temperatures and can, therefore, be sterilized by autoclaving.

Putting safety first, Day Impex announces that its **virucidal disinfectant**, Virkon, is effective against all 17 known virus families, including HIV-1 and hepatitis A or B (*Reader Service No. 133*). This broad-spectrum disinfectant is also an effective bactericide and fungicide. In a single operation, users can disinfect work-surfaces and equipment with Virkon, which is water-soluble, almost odourless, non-corrosive, non-tainting and of a very low toxicity, says Day Impex. Virkon is supplied in powdered form and can be applied directly to spills of blood or body fluids.

According to Life Technologies, cryopreserving a broad spectrum of mammalian cell types need not be a problem with its ready-to-use **cell culture freezing media** (*Reader Service No. 134*). The media are available in two formulations: one with dimethylsulphoxide, the other with glycerol. Both are prepared using Dulbecco's modified Eagle medium. To ensure high viability and sustained cell growth on thawing, fetal bovine and calf sera supplements are used. The media are quality-control tested with both adherent and suspension cells.

Using radioimmunoassay technology, Amersham's healthcare and life science division has introduced a [³H] **scintillation proximity assay (SPA) for platelet activating factor (PAF)** (*Reader Service No. 135*). The assay affords researchers with a reliable and precise means of measuring PAF in biological samples, says Amersham. The SPA technology requires no separation of bound from free liquid and no addition of liquid scintillants. Using a highly specific PAF antibody and a high specific activity tritiated PAF tracer, PAF can be measured in the range of 20–1,280 pg per tube, says Amersham. Each pack contains sufficient reagents for 100 assay tubes. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. To obtain further details, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.